AN INVITATION TO MATHEMATICS

ARNOLD DRESDEN

Robert F. Steinhart

AN INVITATION TO MATHEMATICS

BY

ARNOLD DRESDEN

SWARTHMORE COLLEGE

NEW YORK

HENRY HOLT AND COMPANY

11–45

PRINTED IN THE
UNITED STATES OF AMERICA

DEDICATED TO

PRESIDENT FRANK AYDELOTTE

AND TO THE STUDENTS OF

SWARTHMORE COLLEGE

Now, therefore, go to, proclaim in the ears of the people, saying, Whosoever is fearful and afraid, let him return, and depart early from Mount Gilead: and there returned of the people twenty and two thousand, and there remained ten thousand ("The Book of Judges," chapter VII, verse 3).

The light of the sun gives its greatest joy to the body; the clarity of mathematical verity gives its greatest joy to the spirit (Dmitri Merejkowski, *The Romance of Leonardo da Vinci*, p. 139).

La Mathématique ne vieillit pas; elle ne connaît pas d'hiver. A l'instant où, ployant sous les fruits de l'automne, ses branches pesamment retrouvent la terre nourricière, comme Antée elle reprend des forces, elle rajeunit superbement, et déjà, par l'effet d'une puissante sève circulant dans ces canaux souterrains que nous devons maintenir, pointent les tendres rameaux du printemps (Gaston Julia, *Essai sur le développement de la théorie des fonctions*, 1933).

PREFACE

A few years ago, the faculty of Swarthmore College abolished the requirement of a year's course in mathematics for the A.B. degree. During the discussions which preceded this action, the mathematics department felt called upon to give what justification it could for maintaining the requirement. It had to be conceded that the course usually offered, consisting of some algebra, some trigonometry and some analytical geometry, could hardly be considered adequate for those for whom this course constituted the last formal occupation with mathematics. On the strength of a general conviction, rather than on the basis of actual experience, the statement was made that it would be feasible so to arrange the work for such a course as to open the possibility of gaining some understanding of the significance of mathematics, of its relation to other subjects of study and to human experience. It was suggested to the author that he produce such a course; this was the stimulus which started the work of which this volume is the outcome.

A second incentive was furnished by the frequent inquiries from varied sources for some means of learning why mathematics plays so important a rôle not only in science and engineering, but also in philosophy, why mathematicians so often look upon their work as a form of art, what significance the subject has for the modern world. The most natural answer to such questions was to say that there is no shortcut to the attainment of such understanding, that one desiring to acquire it has to travel a long and laborious road, beginning with the elements of the subject and proceeding as far as his ability will permit; in other words, that such insight is reserved for the professional mathematician. But this reply did not seem to be the only possible one. The feeling persisted that for a person who is willing to make a serious effort, there should be a way of acquiring sufficient comprehension of the nature of mathematics to enable him to understand without having a large technical equipment, the place this subject occupies in the world of thought and that such understanding should form a part of the equipment of an educated person. Thus it appeared worth while to make an attempt at paving the way towards this end.

There is an additional reason why it seemed desirable to the author to do this. It is well known that in many parts of this country there is a growing movement to eliminate mathematics from the list of subjects required for study, not only in the colleges (the action of the Swarthmore College faculty is one example out

of many), but also in the high schools. It is his belief that this movement, in as far as it is guided by sound educational principles, results from the lack of understanding, among educated persons in general and among educational authorities in particular, of the essential character of mathematics. There is little doubt that for this lack of understanding the teachers of the subject are in a large measure responsible. The fact that such a movement can gain adherence after mathematics has been for many years a required subject of study in schools and colleges points to a serious flaw in the manner in which the subject has been presented. There has been too much emphasis on its formal and narrowly technical aspects, even where technique is not the end to be achieved, and neglect of the wider bearings, of the broad human implications of the subject, and of its more interesting and stimulating problems. Perhaps a case can be made for the maintenance of required mathematical instruction in the schools and colleges on the strength of its technical importance alone. But once a fuller appreciation of the subject has been acquired, there is bound to come recognition of the important place it should occupy in the mental equipment of an educated man whether he uses the technique of mathematics or not. Because mathematics has a contribution of fundamental significance to make to the education of our people, we can not allow false conceptions of the subject to persist and to weaken its influence. Mankind may be in a better position to deal with the baffling problems which confront it in the modern world if an understanding of mathematics were the rule rather than the exception.

It is considerations such as these which led the author of this book to undertake a task which he knew to present many difficulties. They determined the end to be sought for, viz., to give a reader who has but little knowledge of the technique of mathematics, an insight into the character of at least some of the important questions with which mathematics is concerned, to acquaint him with some of its methods, to lead him to recognize its intimate relation to human experience and to bring him to an appreciation of its unique beauty. These aims in turn determined to a considerable extent the contents of the book and the method of presentation. From the large and evergrowing domain of mathematics, a few parts had to be selected which involve important ideas and which can be made intelligible without extended prepa-

ration. These parts had to be chosen so as to give some indications of the wide diversity of the field of mathematics and to point to domains in which active work is going on so that the reader might get a sense of the tremendous scope for further development which the subject presents.

Other desiderata to be kept in mind were understanding of the growth of mathematics through the ages, and guidance to sources of further information from which the often meagre treatment to which we had to restrict ourselves could be supplemented. In many instances detailed, complete discussions had to be given up and replaced by descriptive introductions and brief summaries. Frequently proofs have been omitted, so as to lead the reader promptly to a conclusion of which the significance can readily be appreciated and so as to make the result stand out clearly and not too much obscured by details. Incidental remarks of historic or philosophic character have been interspersed wherever an opportunity presented itself.

It is the author's firm conviction that very few people can really learn anything about mathematics by passive reading or by listening to lectures. Active participation is essential. It is for this reason that a considerable number of problems are proposed, over 500 in all. Some of these are intended to give an opportunity of testing whether the subject matter with which they deal has been understood; this testing is especially important for a reader who studies the book by himself. Many others ask for more independent work; they should lead the reader to reflect on the matters with which the book is concerned and to gain some insight into the manner in which mathematics is developed. It is important that the person who works through this material should thoroughly enjoy the experience. The book is not primarily intended as a textbook; much of it is light reading. For this reason it does not try to adhere, either in content or in external form, to the somewhat stereotyped style of the textbook. It is hoped that the reader will have as much pleasure in reading the book as the author has had in writing it.

The material embodied in this volume has been developed in the course of several years with classes at Swarthmore College. Not all the chapters have been used with every group of students. Different sets of chapters can be selected so as to form a satisfactory course for one semester, or for two semesters. If the work succeeds

in arousing the student's interest, he will want to return at a later time to the topics which have been omitted from his course. For the reader who undertakes to study the book without the aid of a teacher, the following indications may be useful. Chapters I, VII and XV can be studied by themselves, although they are not without connections with other chapters. Chapters III–VI present successive developments of the number system and had best be read in that order. There remain three other groups of related chapters, viz., Chapters VIII and IX on Theory of Numbers, Chapters X and XI on Geometry, and Chapters XII–XIV on Analysis.

The preparation that is indispensable for the use of this book does not exceed what is furnished by a good high school course in algebra and in plane geometry. The reader for whom these subjects belong to a rather distant past will do well to supply himself with elementary books in those fields so as to enable him to refresh his memory on the few technical matters which he will be called upon to contribute from time to time. The book expresses the belief that it is not necessary to wait until the graduate school is reached before one can learn something about the more interesting and broadly significant parts of mathematics, that many of the ideas which are involved in advanced parts of the subject can be made accessible to those who bring only a small amount of technical knowledge, provided they also contribute a readiness to concentrate and a taste for abstract thinking.

Whether this belief is sound remains to be seen. In so far as this volume succeeds in its purpose, it puts the belief to a test. In so far as it fails, it will leave the question unsettled. It is hoped that others will improve on the present attempt in order that future generations may make the fundamental and important concepts of mathematics their own more fully than past generations have done.

For valuable advice and helpful criticisms the author is indebted to his friends and colleagues, Drs. H. W. Brinkmann, J. R. Kline and I. J. Schoenberg, who have assisted him in reading the proof sheets. The publishers' interest in this undertaking, their readiness to venture into a new field is a source of hope for the future.

A. D.

Swarthmore, Pa.
June 6, 1936.

CONTENTS

HINTS FOR THE USER OF THIS BOOK

1. References to sections are printed in italics, references to numbered parts of a section are in Roman; e.g., *29*,3 refers to the paragraph numbered 3 in section *29* (page 48).
2. References to numbered equations or to numbered lines of text are enclosed in parentheses, the digits to the left of the decimal point indicate the chapter; e.g., (9.3) refers to the third numbered equation in chapter IX (page 204).
3. At the end of the volume are found (1) an index of definitions, giving the page number on which they are stated; (2) an index of numbered equations giving the page numbers on which they appear for the first time; (3) a page index of the figures; (4) an index of symbols used, giving the page number on which they are explained; (5) an index of theorems and corollaries, giving the number of the page on which they are stated; (6) an index of names and of technical terms. A systematic use of these indices should facilitate the use of the book.
4. A brief commentary on the contents of the book, including discussions of and answers to many of the questions proposed to the reader, is available in a separate pamphlet. It can be secured by application to the publisher.

CHAPTER I

THE STARTING POINT — A FAMILIAR LANDMARK

For the unique value of the Note Book lies in the insight which it affords us in the polarizing quality of a poet's reading — a reading in which the mind moved, like the passing of a magnet, over pages to all seeming as bare of poetic imagination as a parallelogram, and drew and held fixed whatever was susceptible of imaginative transmutation. — John L. Lowes, *The Road to Xanadu*, p. 34.

1. A well-known story and some doubts. It is probably a safe assumption that every one has heard, in seriousness or in jest, of the theorem of Pythagoras. This famous proposition asserts that the square on the hypotenuse (the longest side) of any right triangle is equal to the sum of the squares on the other two sides. Legend has it, that when Pythagoras discovered it "he offered a splendid sacrifice of oxen." We shall make it the starting point for a few excursions in the field of mathematics.

We shall not be concerned with a proof of this beautiful theorem. In the course of the centuries, many proofs have been obtained; any textbook on elementary geometry supplies at least one of them. It is still a matter of conjecture just how Pythagoras proved it. There even seems to be considerable doubt whether the honor of having discovered it belongs to Pythagoras.[1]

There are however several aspects of the theorem that we shall find it worth while to consider. In the first place, we make the simple but important observation, that its validity is independent of time and place. Whether I draw a triangle ABC here on this paper and then construct squares on each of its sides; or whether Julius Caesar had made such a construction on the shores of the Rubicon, or whether the Grand Lama of Tibet performs a similar operation, the same conclusion is always reached: the sum of the squares on the sides of the right angle is equal to the square on the hypotenuse. This simple statement of Pythagoras has a scope

[1] See, e.g., T. L. Heath, *A Manual of Greek Mathematics*, pp. 95–100. A large collection of proofs of the theorem is found in the interesting little book by W. Lietzmann, *Der Pythagoreische Lehrsatz* (B. G. Teubner, Leipzig).

which sweeps over the ages and over all space. This it has in common with most mathematical theorems.

On the other hand, an element of doubt may well enter our minds. Not whether the proposition of Pythagoras might ever fail to be true, but whether any one, Julius Caesar, or the Grand Lama of Tibet, or the Grand Duke Josephus, *can* ever draw a right triangle; whether we can construct squares on the sides of triangles we can't draw, and whether we can determine by measurement the areas of squares we can't construct.

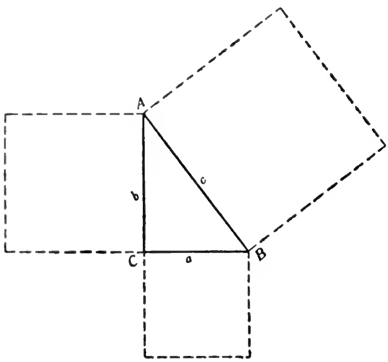


FIG. 1

With regard to each of those questions there may well be considerable uncertainty. We are therefore led to ask what then the proposition is really about. The answer would be: it is about right triangles. Very well — but what is a right triangle? Has any one ever met one in the street? Can we put our hands on one? To these last questions the answer has to be in the negative. We have to say in fact that the proposition is not concerned with any concrete object. We can make pencil marks which create an illusion of right triangles and squares, but they only serve as reminders, as tangible objects of which the triangles and squares are the shadows. The Pythagorean theorem deals with abstractions. This also it has in common with most results obtained in mathematics.

2. From lines to numbers. We are now going to consider a different aspect of this proposition, one more properly concerned with it alone. It is well known from elementary geometry that if the length of a line be p inches, then the number of square inches contained in a square constructed on that line as a side is equal to $p \times p$; and from elementary algebra we recall that $p \times p$ is denoted by the symbol p^2 (read "p square," or "p to the second power").[1]

[1] If the line, measured in a different linear unit, in centimeters, or in feet, or in yards, had a measurement q, then the corresponding square would contain q^2 square units, each of which is a square whose four sides are each equal to one of these linear units, i.e. q^2 square centimeters, q^2 square feet, or q^2 square yards.

We can say therefore, no matter what unit is used for the measurements, provided it be the same for all three sides: if the measures of the sides of the right angle in our right triangle ABC are a and b, and that of the hypotenuse c (see Fig. 1), then the theorem of Pythagoras asserts that

$$a^2 + b^2 = c^2. \tag{1.1}$$

This is the algebraic form of the theorem.

It is now interesting to inquire whether it is possible to construct a right triangle, such that the measurement of each of its sides is a whole number. In non-geometrical language, the question is, whether we can find three positive *integers* (*integer* means "whole number") a, b, c, such that the relation (1.1) holds between them. Doubtless every reader of this chapter is acquainted with one answer to the question. For the simple fact that $9 + 16 = 25$ can at once be put in the form $3^2 + 4^2 = 5^2$, from which we conclude that $a = 3$, $b = 4$, $c = 5$ is such an answer. We know then that if on the two sides of a right angle segments are laid off equal to 3 and 4 linear units respectively, the length of the line joining their extremities will be 5 of these units.

This answer having been obtained a number of other questions now suggest themselves. The reader will notice that the road we have followed so far leads us from a theorem in geometry to questions concerning numbers; clearly there can not be any formidable barrier between geometry and number theory — the only tariff to be paid is that of concentrated attention. But, on with our questions:

1. Do there exist other sets of three positive integers a, b and c, which are so related that the relation (1.1) is verified?
2. If such triples exist, how can they be found?
3. Is there among them any other set of three consecutive integers, besides 3, 4 and 5?

3. Clearing the ground. The first of these questions can readily be answered in the affirmative; for instance, since $25 + 144 = 169$, we recognize at once that $a = 5$, $b = 12$, $c = 13$ is a solution of equation (1.1). It will be worth while to try to find by a hit-or-miss method two or three more such sets (see *4*).

Before engaging in this quest, we must be clear on one or two points: the first of these is when we are to consider two different triples of numbers as furnishing different solutions of equation

(1.1). But to begin with, we must have a convenient and brief form in which to express our statements. We shall use the phrase "$(p, q; r)$ satisfies (1.1)," or "$(p, q; r)$ is a solution of (1.1)," or "$(p, q; r)$ is a Pythagorean triple" as equivalent to the statement "the numbers p, q and r are such that $p^2 + q^2 = r^2$." It follows at once that of $(p, q; r)$ and $(q, p; r)$ either is a Pythagorean triple if the other is.

The sort of question we have to consider is the following: Are we to look upon (6, 8; 10) as being a solution different from (3, 4; 5)? If so, only the supply of writing materials and of time would limit the number of Pythagorean triples we could write down, as soon as any one is known. For if $(p, q; r)$ satisfies (1.1) then $(kp, kq; kr)$ will satisfy that equation as well, no matter what integer be taken for k. To distinguish such pairs of solutions from a pair of solutions like (3, 4; 5) and (5, 12; 13), we shall say that $(p, q; r)$ and $(kp, kq; kr)$ are a pair of dependent solutions, while (3, 4; 5) and (5, 12; 13) are a pair of *independent* solutions. The interesting part of our problem is then to find independent solutions of (1.1), i.e. solutions no two of which are a dependent pair.

Two further remarks are in order here:

(a) Of two dependent solutions, the one which involves the smaller numbers will be the most convenient to deal with. If we agree then to use of any set of dependent solutions the one involving the smaller numbers, the last statement in the preceding paragraph amounts simply to this, that we want to find such solutions of (1.1) as consist of three integers which do not have a factor in common. Let us agree to call such a solution a *primitive* solution. We would, instead of the non-primitive solution (10, 24; 26) consider the primitive solution (5, 12; 13). To abbreviate, we shall use p.P.t. to indicate "primitive Pythagorean triple."

(b) We have been asking for solutions of (1.1) which consist of positive integers. From the discussion just preceding, we conclude that from any primitive solution in positive integers $(p, q; r)$ we can obtain also an unlimited number of solutions in positive rational numbers. (For our present purpose we shall understand the term "rational number" to mean "integer, or common fraction"; when we speak about "real numbers" before we have given a sharp and clear definition of them, we shall mean any number that can be conceived of as representing a simple physical measurement.) For, the statement that $(kp, kq; kr)$ is a solution whenever $(p, q; r)$ is

remains true if k, instead of being an integer, is a common fraction. Thus from the primitive solution (3, 4; 5) we obtain at once the solutions $(\frac{3}{2}, 2;\ \frac{5}{2})$, $(2, \frac{8}{3};\ \frac{10}{3})$ and many others. But, conversely, to every set of positive rational numbers that satisfies equation (1.1), there corresponds a set of integers which is a primitive Pythagorean triple. For instance, since

$$\left(\frac{25}{3}\right)^2 = \frac{625}{9},\ 8^2 = 64,\ \left(\frac{7}{3}\right)^2 = \frac{49}{9},\ \text{and}\ \frac{49}{9} + 64 = \frac{49+576}{9} = \frac{625}{9},$$

we see that $(\frac{7}{3}, 8;\ \frac{25}{3})$ is a solution of (1.1), i.e. that

$$(\tfrac{7}{3})^2 + 8^2 = (\tfrac{25}{3})^2;$$

but this statement is equivalent to the statement $7^2 + 24^2 = 25^2$ (why?), i.e. that (7, 24; 25) also satisfies (1.1).

And, in general, suppose that $\left(\frac{p_1}{q_1}, \frac{p_2}{q_2};\ \frac{p_3}{q_3}\right)$ is a Pythagorean triple, and that $p_1, q_1;\ p_2, q_2;\ p_3, q_3$ are positive integers. Then

$$\left(\frac{p_1}{q_1}\right)^2 + \left(\frac{p_2}{q_2}\right)^2 = \left(\frac{p_3}{q_3}\right)^2.$$

But from this it follows that

$$(p_1q_2q_3)^2 + (q_1p_2q_3)^2 = (q_1q_2p_3)^2,$$

which means that $(p_1q_2q_3,\ q_1p_2q_3;\ q_1q_2p_3)$ is a solution in integers. If this is not a primitive Pythagorean triple, we can deduce one from it.

In the numerical example just considered, we had $p_1 = 7$, $q_1 = 3$; $p_2 = 8$, $q_2 = 1$; $p_3 = 25$, $q_3 = 3$. Hence $p_1q_2q_3 = 21$, $q_1p_2q_3 = 72$, $q_1q_2p_3 = 75$. From the solution in rational numbers $(\frac{7}{3}, 8;\ \frac{25}{3})$, we obtain therefore by this method the non-primitive solution in integers (21, 72; 75); this in turn leads to the primitive solution (7, 24; 25).

The upshot of this discussion is then that the problem of finding all solutions in positive rational numbers of our equation (1.1) is equivalent to that of finding all the positive primitive Pythagorean triples, in this sense that every solution in positive rational numbers of the equation (1.1) can be obtained from a positive p.P.t. by multiplying or dividing its numbers by a positive integer; e.g. the solution $(\frac{7}{3}, 8;\ \frac{25}{3})$ is obtained from (7, 24; 25) upon division of its numbers by 3.

We have here a simple illustration of the fact that, even though one statement be in form and in words more general than another, its content and real significance may not be any greater. The problem of finding solutions of equation (1.1) which consist of positive rational numbers has been reduced to that of determining the Pythagorean triples, which consist of positive integers without common factor. We are now ready to proceed with the questions stated at the end of *2*. But it will be desirable to have a short recreation period before we continue our journey.

4. First resting place.

1. Hunt for 3 sets of 3 positive integers each, for which equation (1.1) is satisfied (see p. 3).

2. Devise a geometrical argument to show that *either* $(p, q; r)$ and $(kp, kq; kr)$ are both solutions of equation (1.1), or *else neither* of them is; k is understood to be any rational number.

3. Show that there is no gain in generality if we admit negative numbers in the solution of equation (1.1).

4. Show that $(3, 4; 5)$ is the *only* solution of equation (1.1) which consists of *consecutive* positive integers.

Hint. Three consecutive positive integers can be represented in the form

$$n - 1, n, n + 1.$$

5. Determine *all* solutions of equation (1.1) which consist of *consecutive* integers.

6. Three numbers p, q, r can be arranged in *6* different ways. Can it ever happen that these three positive integers p, q, r form a Pythagorean triple if arranged in any of two or more different ways? Explicitly stated: are there any positive integers p, q and r such that any two or more of the sets $(p, q; r)$, $(p, r; q)$, $(q, p; r)$, $(q, r; p)$, $(r, p; q)$, $(r, q; p)$ are Pythagorean triples?

7. Determine all Pythagorean triples which consist of terms of an arithmetic progression. (An arithmetic progression is a sequence of numbers for which the difference between two successive numbers is constant; e.g., . . . 5, 12, 19, 26, . . .; or . . . -12, -16, -20, . . .).

8. Prove that no Pythagorean triple exists in which one number is a mean proportional between the other two. (To say that "b is a mean proportional between a and c" means that $ac = b^2$).

9. Determine the p.P.t. that can be obtained from the Pythagorean triple $(\frac{32}{5}, 51; 51\frac{2}{5})$.

10. Show that $(2n + 1, 2n^2 + 2n; 2n^2 + 2n + 1)$ is a Pythagorean triple for every value of n.

11. Do any Pythagorean triples exist of the form $(p, q; r)$ in which $r = q + 1$, except those of the form given in 10?

12. Show that $(2m^2 + 2mn, 2m^2 + 2mn + 1; 2m^2 + 2mn + n^2)$ is a Pythagorean triple, provided m and n satisfy the relation $2m^2 = n^2 - 1$; show that the statement is still true if the sign of the terms involving mn is changed.

5. The first ascent. The exercises in *4* having been completed, we have obtained the answers to the 1st and 3rd questions of *3;* moreover we should be well prepared for dealing with the 2nd question, viz. with the problem of finding all p.P.t.'s. To attack it we shall gradually narrow down the class of numbers from which they may be selected. The first step is the following:

Lemma 1.[1] In a primitive Pythagorean triple no two numbers have a common factor.

Proof. Suppose that $(p, q; r)$ is a p.P.t. and that p and r had the integral factor f in common; then we could put $p = fp_1$ and $r = fr_1$, in which p_1 and r_1 are integers. The following relations would then result:

$$f^2p_1^2 + q^2 = f^2r_1^2,\ q^2 = f^2r_1^2 - f^2p_1^2 = f^2\ (r_1^2 - p_1^2).$$

We would conclude from this last statement that $q = f\sqrt{r_1^2 - p_1^2}$, from which it would follow, as any child can understand, that q had the factor f.[2] But then p, q and r would have the common factor f; however, this is excluded by the hypothesis that $(p, q; r)$ is a *primitive* Pythagorean triple.

The reader should have no difficulty in completing the proof by showing that neither p and q, nor q and r can have a factor in common (see *6*, 2).

We shall frequently be concerned with a set of two or more integers, which have no common integral factor; also with single integers which have no integral factors besides themselves and unity. For convenience of reference we shall state explicitly the definitions of the terms which are used to designate such integers.

Definition I. Two or more integers which have no integral factor in common besides 1 are called *relatively prime.*

[1] A lemma is an "auxiliary proposition" — see Heath, *A Manual of Greek Mathematics*, p. 216.

[2] Evidently $r_1^2 - p_1^2$ is an integer; moreover, being equal to $\frac{q^2}{f^2}$, it is a perfect square. Therefore $\sqrt{r_1^2 - p_1^2}$ represents an integer.

For example, 10 and 9; 36 and 55; 35, 87 and 121 are sets of relatively prime integers.

Definition II. An integer which is different from 1 and has no integral factor except itself and 1 is called a *prime number;* an integer which is different from 1 and which is not a prime number is called a *composite number.*

The reader will have little difficulty in locating the first 10 prime numbers. But there are a great many simple-sounding questions concerning prime numbers which he would have real difficulty in answering. Among them,[1] we will only mention in passing the problem of determining the number of prime numbers between two given numbers. The alluring side road which invites us here would soon lead us to formidable barriers; great strength and much insight are needed for progress. We return to our simple task.

From Lemma 1 it follows that no p.P.t. can contain two even numbers. As a partial counterpart, we will now prove the following:

Lemma 2. If $(p, q; r)$ is a Pythagorean triple, then p and q can not both be odd.

Proof. Every even number is divisible by 2; hence every positive even number must have the form $2n$, where n is a positive integer. But then every odd positive number must have the form $2n + 1$, where n is a positive integer or 0.

Suppose now that p and q are both *odd* positive integers; we can then put

$$p = 2p_1 + 1 \quad \text{and} \quad q = 2q_1 + 1,$$

in which p_1 and q_1 are integers. If then $(p, q; r)$ is a Pythagorean triple, we have

$$r^2 = p^2 + q^2 = (2p_1 + 1)^2 + (2q_1 + 1)^2 = 4(p_1^2 + q_1^2 + p_1 + q_1) + 2.$$

The number represented by the right-hand side of this equation will leave a remainder 2, if it is divided by 4; hence it is divisible by 2, but not by 4. This means that it has *only one* factor 2, so that it can not be the square of an integer. Consequently, the hypothesis that p and q are both odd is in contradiction with the hypothesis that $(p, q; r)$ is a Pythagorean triple.

Corollary. If $(p, q; r)$ is a p.P.t., then of the two numbers p and q, one must be even and the other odd.

[1] See, e.g., Carmichael, *The Theory of Numbers*, pp. 28–9.

This corollary is an immediate consequence of Lemmas 1 and 2.

Lemma 3. If t and s are positive integers ($t > s$),[1] such that $t + s$ and $t - s$ have a common factor k, then t and s have in common all the factors of k, except possibly a single factor 2.

Proof. Suppose $t + s = ku$, and $t - s = kv$. Then $2t = k(u + v)$ and $2s = k(u - v)$. Therefore all the factors of k must occur in $2t$ *and* in $2s$. It follows that all the factors of k, except possibly a single factor 2, are common to t and s.[2]

The lemmas which have been proved will now enable us to determine all primitive Pythagorean triples. Suppose that $(p, q;\ r)$ is such a set of numbers. In accordance with the corollary on page 8, we may suppose that q is odd and p even; it follows then from Lemma 1, that r is also odd. Moreover, since

$$p^2 + q^2 = r^2,\ p^2 = r^2 - q^2 = (r + q)(r - q).$$

Now, since r and q are both odd, $r + q$ and $r - q$ are both even, i.e. they each have a factor 2. But moreover they have no other factor in common; for, if they did, r and q would then also have that factor in common (Lemma 3), which would contradict Lemma 1. If then the product of $r + q$ and $r - q$ is to be a square of an integer, there must exist two integers u and v, without common factors, such that

$$r + q = 2u^2,\ r - q = 2v^2 \quad \text{and} \quad p^2 = (r + q)(r - q) = 4u^2v^2.$$

Consequently, $r = u^2 + v^2$, $q = u^2 - v^2$, $p = 2uv$.

Clearly, if r and q are to be odd, u and v must not only be relatively prime, but neither can they both be odd. We conclude that if $(p, q;\ r)$ is a p.P.t., there must exist two integers u and v, $u > v$, relatively prime and not both odd, such that

$$p = 2uv,\ q = u^2 - v^2 \quad \text{and} \quad r = u^2 + v^2. \tag{1.2}$$

But conversely, if p, q and r are determined from (1.2), in which u and v are two relatively prime integers not both odd, then $(p, q;\ r)$ is a primitive Pythagorean triple. That it is a Pythagorean triple, is readily verified; for $(2uv)^2 + (u^2 - v^2)^2 = 4u^2v^2 + u^4 - 2u^2v^2 + v^4 = u^4 + 2u^2v^2 + v^4 = (u^2 + v^2)^2$. That it is a primitive triple can easily be seen; the proof can be put in words more conveniently, if we introduce the word "parity." Two

[1] The reader will recall that the symbol > stands for "greater than," and < for "less than."

[2] Compare p. 136.

integers will be said to have equal *parity* if they are both even *or* both odd; they will be said to have unequal *parity*, if they are neither both even nor both odd, i.e. if one is even, the other odd. Since u and v have unequal parity, u^2 and v^2 are of unequal parity, and hence q and r are odd; consequently p, q and r can not have an even factor in common. Suppose then q and r had an odd factor in common. Since $2u^2 = r + q$ and $2v^2 = r - q$, u^2 and v^2 would each have this factor; but this can not be the case if u and v are relatively prime.

The question with which we started has thus been answered completely. For convenience of reference, and also in order to give the result some prominence, we record the answer we have found.

Theorem I. Whenever u and v are relatively prime integers of different parity, and $u > v$, the numbers

$$p = 2uv,\ q = u^2 - v^2,\ r = u^2 + v^2$$

form a primitive Pythagorean triple $(p, q;\ r)$; moreover, every primitive Pythagorean triple is of that form.[1]

We shall have a good bit more to say both about this theorem and about its proof in later chapters.[2]

6. Rest after the climb.

1. Determine all p.P.t.'s which involve positive integers less than 300.

2. Prove that if p and q, or q and r have a common factor, then $(p, q;\ r)$ can not be a p.P.t. (Compare p. 7.)

3. Prove that every even integer which is a perfect square has the factor 4.

4. Prove that if t and s $(t > s)$ are positive integers such that $2t + s$ and $t + 2s$ have a common factor, then t and s must also have this factor in common, except possibly for a factor 3.

5. What can be said about t and s, if it is known that $2t - s$, and $t + 2s$ have a common factor?

6. The questions in 4 and 5 are special instances of a more general question. Formulate this question; then answer it.

7. Use Theorem I to obtain a new proof of the statement made in *4*, 4 (p. 6).

[1] It should be clear to the reader that we might equally well put $p = u^2 - v^2$ and $q = 2uv$; this form is in closer accord with the examples of a p.P.t. which have occurred thus far, since in them the second element is an even number.

[2] A simple discussion and proof of this theorem is found in the book by Lietzmann, *op. cit.*

8. Prove that there exists no isosceles right triangle whose three sides are measured by integers.

9. Show that if integers x, y and z satisfy the equation $2x^2 + y^2 = 3z^2$, and if x and y have a factor f in common, then f is also a factor of z.

10. What can be said of y in 9, if x and z have a factor in common? and what about x, if y and z have a factor in common?

11. Discuss common factors of integers x, y and z which satisfy the equation

$$3x^2 + 4y^2 = 5z^2.$$

12. Show that no integer of the form $4n + 3$ can be a perfect square.

7. Brief retrospect. We have now completed our study of the equation (1.1) which is connected with the familiar theorem of Pythagoras. In the following chapter we shall pick up some of the threads which have gone into the woof and warp of its design. They will lead us to regions perhaps surprising, but at any rate full of interest, scientific, philosophic and, I think, poetic.

Let us recall at this point that the theorem of Pythagoras has led us to Theorem I, and that incidentally we have become acquainted with the important concepts of "prime numbers" and of "pairs of relatively prime integers."

CHAPTER II

THE FIRST EXCURSION — NEW VIEWS FROM AN OLD TRAIL

To have learned to open the mind to hitherto unknown and even inconceivable states of thought and feelings, is to have undergone a permanent change. It is like learning to swim. Once a swimmer always a swimmer. — Alfred E. Zimmern, *Learning and Leadership*, p. 51.

We begin by inquiring a little further into the full meaning and significance of Theorem I. At the beginning of Chapter I, we called attention to the fact that the validity of the theorem of Pythagoras was independent of time and place. We are concerned now not with the scope of the validity of our Theorem I, but with the scope of its content. With a slight change in language, Theorem I tells us that (1.2) gives Pythagorean triples for *all* relatively prime pairs of integers u, v of unequal parity ($u > v$), and that *all* p.P.t.'s are obtainable from such pairs of integers. We shall pick out one word in this statement as the starting point for our present excursion.

8. What is " all "? We observe that the word *all* as used in the preceding statement of Theorem I has a sense quite different from that usually attributed to it in ordinary language. When we speak, for instance, of "all human beings," "all houses in a town," "all the people in a room," we have in mind a definite totality; at least a totality which is definite at any one time and which could be counted off. It may very well be that the accurate counting of such totalities is a process hard to carry out in actual practice, particularly if the totality should be subject to change while the counting process goes on. It is a notorious fact for instance, that practically every census count is out of date before its results are available to the public. In spite of this, we *can conceive* of schemes whereby such totalities as are designated by the use of the word "all" in ordinary language, could be counted. It may be an interesting game for the reader to try to work out such schemes for totalities like those mentioned above. He may have to have

recourse to fantastic devices, particularly if his ambition should extend to "all the trees on the surface of the earth," "all the leaves on all the trees in Delaware County." Such speculations have a peculiar fascination, but they are a bit too far removed from our present purpose to justify us in following them further. Let us "return to our mutton."

When we speak of "*all* primitive Pythagorean triples," or of "*all* such pairs of integers u and v," there appears a difficulty of quite a different character. For, no conceivable process would enable us to count the positive integers — let alone the p.P.t.'s. If any person should start counting integers the moment he is born, and were condemned to continue counting throughout his life, he would not come to the end, no matter how rapidly he could count. If Adam had begun counting integers the moment he appeared in the garden of Eden and if every successive generation since Adam had carried this counting forward up to the present time, we would not have exhausted the totality of integers — nay, we would be, in a certain real sense, as far from the end as Adam was. If a counting machine were devised which would click every time a leaf bud opened on any tree anywhere on the surface of the earth and if this machine had been in operation for two million years, the total number of clicks this machine had made would be as nothing if compared to the totality of p.P.t.'s.

The long and short of these remarks is simply this, that the positive integers, the p.P.t.'s, "such pairs u, v," form totalities in which we can count a first element, but no last element; while the other totalities which have been introduced for the purpose of comparison all possess not only a first element but also a last element, however far that last element may be removed from the first in the counting process. Collections which have no last element, attainable by counting, are of frequent occurrence and of great importance.

We shall therefore introduce a special name to designate them.

Definition III. A collection in which the counting process can not be carried on to a last element, is called an *infinite set*. By contrast, totalities which can be completely counted off from a first element to a last element will be called *finite sets*.

9. Infinite sets — and a moral. With this definition in mind, it is evident that the word "all" as used in Theorem I, where we speak of "all primitive Pythagorean triples," "all such pairs u, v"

etc., refers to infinite sets. Not only these totalities but also "all positive rational numbers," "all positive integers," "all the points on a line segment," "all the common fractions between 0 and 1" are infinite sets. The reader will do well to familiarize himself with the concept of an infinite set by thinking up further examples. Having accumulated a sufficient store of concrete instances to make the abstract concept significant, we can proceed to a further study of these collections, to which this consideration of Theorem I has led us. The first question one would naturally ask is this: Can numbers be attached to infinite sets as they can to finite sets?

To prepare the answer to this question we might inquire how numbers are attached to finite sets — and the most probable answer to this query would be "by counting them off." If this were all that could be said we would not have gained anything, for we have already seen that infinite sets can not be counted off. But, if we go back to a primitive form of counting we shall discover an essential characteristic of the process and therein a method which is applicable also to infinite sets. When an unsophisticated barbarian, whether on the slopes of the Himalaya, in the Australian bush, or in New York City, wants to count a small collection of things, he "counts them off on his fingers," that is to say, he pairs off the objects he wants to count, one by one, with his fingers. Or, he makes a notch in a stick for each object, or a mark on a wall thus pairing off his objects with the notches or the marks.[1]

Thus we recognize as an essential element in "counting a collection of things," the pairing off of the things in the collection with those of another set of things, be they fingers and toes, or marks on a stick, or the marks which we now call number signs. And the essential aim pursued in "counting a collection of things" is not to obtain a number but rather comparing the collection with another collection, already well known like the set of fingers or readily seen as a whole like notches in a stick.

And this aim can be accomplished without the use of numbers by direct comparison. At a properly arranged dinner party, there is a chair for every person and there are no vacant chairs. To every chair there corresponds a guest and to every guest a chair. There

[1] Out of such marks, number symbols have been developed in the course of centuries. We get here a glimpse into a region which is very interesting on its own merits, but which we must pass by. See e.g., T. Dantzig, *Number, the Language of Science*, 2nd edition, Chapters I, V; L. Conant, *The Number Concept*.

is a *one-to-one* (usually written 1–1) *correspondence* between chairs and guests. If the hostess wants to be certain that there are enough chairs and not too many she can *either* count chairs and guests and make sure that she has the same number for each, *or* she can pair off guests and chairs, i.e. set up a 1–1 correspondence between chairs and guests. If the dinner party were magnified into an international-rotary-clubs banquet, either of the two methods could still be used. But if it were to include "all" the offspring of an interminably prolific race of march-hares, the first method would certainly no longer be available, because this offspring could not be counted. But the second method remains applicable even in this case, provided we can set up a 1–1 correspondence between the march-hares and the chairs. Thus we understand that the essential aim of counting, viz. that of comparing collections can be carried over to infinite sets by the use of 1–1 correspondences. Such a correspondence clearly can not be set up for infinite sets by pairing off element for element; instead there is required a rule which enables us to associate with *any given* element of one set, a definite element of the other set. Having thus cleared the ground let us lay down some of these notions in definitions.

Definition IV. A one-to-one (1–1) correspondence between two collections A and B is established whenever there is a rule which pairs with each element of A one and only one element of B, and with each element of B, one and only one element of A.

Definition V. Two sets are called *equivalent* if a 1–1 correspondence exists between them.

Remark. For finite sets this concept of "equivalence" reduces to what is ordinarily called "equality in number."

Infinite sets have no "number," but the concept of equivalence applies to them equally as well as to finite sets. We have thus extended to infinite sets the solution of one of the problems which, in the case of finite sets, can be accomplished by means of counting, viz. to decide whether two finite sets are equal in number. It remains to deal with the extension to infinite sets of some other uses of counting, viz. of the processes underlying the concepts "less than" and "greater than" as applied to finite sets. But some examples and exercises are long overdue.

(a) The simplest infinite set is the collection of the positive integers, or rather better, the collection of the natural numbers

1, 2, 3, . . . to which neither a positive nor a negative sign is attached.[1] The even integers (i.e. the multiples of 2 or numbers of the form $2n$), the odd integers, i.e. the numbers of the form $2n + 1$, the integers of the form $3n$ (i.e. multiples of 3), those of the form $3n + 1$, or $3n - 1$, or $4n + 1$ are other simple examples of infinite sets.

A fundamental property of infinite sets is exhibited by the fact that the set of natural numbers is equivalent to that of the even integers. In accordance with Definition V, the proof of this statement requires that a rule be set up, which pairs a single even integer with every integer, *and* a single integer with every even integer. This pairing off is accomplished very easily; it is suggested by the following arrangement:

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & \ldots \\ 2 & 4 & 6 & 8 & 10 & 12 & \ldots \end{array}$$

and consists in pairing off 2 with 1, 4 with 2, and so forth, in general, the even integer $2k$ with the integer k, and conversely. This rule is applicable to *every* element in each set, even though the sets are infinite; exactly because they are infinite, such a rule is needed to set up effectively a 1–1 correspondence between these sets.

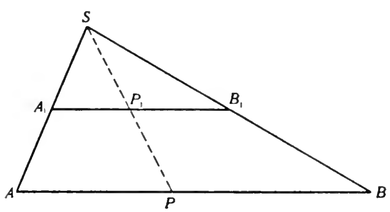


FIG. 2

(b) An infinite set of a more complicated sort is the collection of points on a line segment, 1 inch long. Another illustration of the property of infinite sets, alluded to above, is found in the fact that the set of points on a line segment, 1 inch long, is equivalent to the set of points on a line segment, $\frac{1}{2}$ inch long. The proof can be made, in a variety of ways, all essentially the same. Let us suppose, see Fig. 2, that the lines AB and A_1B_1 are parallel and that the lengths of AB and A_1B_1 are 1 inch and $\frac{1}{2}$ inch respectively. We draw the lines AA_1 and BB_1, and produce them to

[1] A study of the relation between the natural numbers 1, 2, 3, . . . without sign, on the one hand and the set of positive integers on the other hand is well worth while. The interested reader will find a good starting point for such a study in B. Russell, *Introduction to Mathematical Philosophy*, p. 63. See also *39*, 1 (p. 69).

their meeting point S (will the lines AA_1 and BB_1 surely meet?). A 1–1 correspondence between the points of AB and those of A_1B_1 is now set up if every point P on AB is paired off with that point P_1 in which A_1B_1 is met by SP.

These examples should suffice to point out to the reader the significance of the definition of "equivalence of sets." In particular they should make it clear to him that what is considered as common sense when we deal with finite sets, is no longer common sense in the realm of infinite sets. When we enlarge our universe so as to include infinite sets, we have to modify our common-sense notions; for some of these notions are *not common* to all the inhabitants of the extended universe. There is a useful hint contained in this phenomenon; it should lead us to inquire with regard to accepted common-sense notions, how far their scope extends, to what domains they are applicable, and to be prepared for considerable modifications of their significance when we enlarge or restrict the universe of discourse. What is looked upon as common sense in relations between nations, looked at from a national point of view, may turn out to be uncommon nonsense when looked at from an international point of view. Many more useful analogies could be drawn; but — back to our mutton.

10. Learning to swim.

1. Show that the following pairs of infinite sets are equivalent:

(a) the natural numbers and the natural numbers of the form $3n + 1$;

(b) the natural numbers and the natural numbers of the form $5n - 2$.

2. Prove that if the sets A and B are equivalent, and also the sets B and C, then the sets A and C are equivalent; in particular, verify that A is equivalent to itself.

3. Show that the set of natural numbers is equivalent to the set of integers 7, 8, 9, . . .

4. Prove that if an infinite set B is obtainable from a set A by omission of a finite number of elements, then A and B are equivalent sets.

5. Is it possible for a set B to be equivalent to a set A, if it is obtained from the latter by the omission of an infinite number of elements?

6. Show that the set of points on the circumference of a circle whose radius is 1 inch is equivalent to the set of points on a line of 1 inch length.

7. Prove the equivalence of the two sets consisting of the points on two segments of the same line, each 1 inch long.

8. Prove that the set consisting of the eyes of all non-defective mammals is equivalent to the set of their ears; also that the set consisting of their noses is equivalent to the set consisting of their pairs of forelimbs.

9. The set A consists of the natural numbers; the set B of all pairs of natural numbers of the form $(n, n+1)$, such as $(5, 6)$, etc. Prove that the sets A and B are equivalent.

10. The set B consists of all pairs of natural numbers, having the forms $(n, n+1)$, $(n, n+2)$ or $(n, n+3)$; the set A is the set of natural numbers. Prove the equivalence of A and B.

11. Prove that the set of proper common fractions [1] whose numerator is 1, 2, 3, 4 or 5 less than the denominator is equivalent to the set of natural numbers.

12. Prove that the set of proper common fractions is equivalent to the set of common fractions whose value lies between 1 and 2; or between 2 and 3.

11. Old words with a new meaning. The reader who has worked through the exercises in *10*, will have noticed that there are many sets equivalent to the set of natural numbers. Such sets are designated by a special name.

Definition VI. A set which is equivalent to the set of natural numbers is called a *denumerable set.*[2]

It would be natural to inquire whether there are any infinite sets which are not denumerable. The answer is: yes, there are, the set of all points on a line segment is not denumerable. The justification for this answer can not very conveniently be given until we have studied the comparison of infinite sets; it is therefore postponed to the end of the chapter (see p. 29).

The comparison of non-equivalent infinite sets is based upon the concept "part of a set," defined as follows:

Definition VII. If all the elements of a set B belong to a set A, then B is said to be a *part* of A; if, moreover, there are elements of A which do not belong to B, then B is said to be a *proper part* of A.

It follows from this definition that every set is a "part," but not a "proper part" of itself. The set consisting of the human beings in Pennsylvania is a part of the set consisting of the human

[1] A proper common fraction is a fraction of the form $\frac{a}{b}$ in which a and b are natural numbers such that $a < b$.

[2] Instead of "denumerable" the words "enumerable" or "countable" are frequently used. The latter term is perhaps confusing because it suggests the possibility of counting off such sets; and this we have seen to be impossible.

beings in Pennsylvania and a proper part of the set consisting of the human beings on the Continent of North America; the set of even numbers is a proper part of the set of natural numbers; the set of all anthropoids is a part of the set of all mammals. Is the set of all fishes a part of the set of mammals? the set of all onions a part of the set of all lilies?

We are now able to define the relations of inequality between sets.

Definition VIII. If the set A is equivalent to a proper part of the set B, and B is not equivalent to a proper part of A, then A is said to be *smaller* than B, or B *greater* than A.

The definitions of "part," "smaller," "greater," will probably seem unnecessarily complicated to the reader imbued with the common sense of finite sets. For, he will ask, if A is equivalent to a proper part of B (i.e. if, after removing some elements from B, there is 1–1 correspondence between what is left and the whole of A), how could, at the same time, a proper part of A be equivalent to the whole of B? And it is clear that with finite sets this could not occur. But if we are dealing with infinite sets, there is nothing surprising in such an apparent contradiction of common sense. Indeed, it is easy to give an example in which it occurs; we have but to consider the set A of natural numbers and the set B of the even integers. We have already observed that B is a proper part of A; hence, we can say, since every set is equivalent to itself (see *10*, 2; p. 17), that B is equivalent to a proper part of A. But on the other hand the sets 1, 2, 3, 4, . . . and 4, 8, 12, 16, . . . are also equivalent, while the latter is a proper part of B; hence A is also equivalent to a proper part of B.

The example we have just considered is a special case of a general theorem, which has been proved only a comparatively short time ago, viz.

If a set A is equivalent to a part of the set B, and B is equivalent to a part of A, then A and B are equivalent sets.

Unfortunately, the proof of this beautiful theorem known as the *equivalence theorem* requires more preparation than it is possible to give here. The theorem must be for us like a lovely castle that we can see clearly from the road we are traveling but that we can not reach to explore its hidden treasures unless we climb an arduous path. If the climb does not fit in with our other plans, we journey onwards casting a longing look on the castle and promising ourselves some day to undertake the voyage to its gates.

The equivalence theorem is of considerable importance in the general theory of infinite sets, to which the present chapter is an introduction. We shall stop our further developments for a moment in order to glance at the history of this theory and to get somewhat oriented in the literature of this subject.

12. A little history. The theory of infinite sets was founded by *Georg Cantor* (1845–1918) in the years from 1870 to 1890. Cantor was professor of mathematics at the University of Halle, when his first contribution to this theory was published, in 1874, in volume 77 of the celebrated *Journal für die reine und angewandte Mathematik.* Not many of the leaders of mathematical thought at that time received his revolutionary-sounding ideas with approval; not many of them were sufficiently free from obsession by "common-sense" notions to enable them to accept Cantor's ideas. A good many years passed by before Cantor's theory was generally accepted among mathematicians. Since the beginning of the present century however, it has become an essential tool in many important parts of mathematics. But although the objections which it met in its early stages have largely been overcome it has given rise to a new crop of doubts. The theory which even in English is frequently referred to by its German name "Mengenlehre" (the doctrine of sets) has led to questions which affect in a very fundamental way the philosophical bases of mathematics. A considerable literature has grown up dealing with the paradoxes of "Mengenlehre." Within the last 15 years there has been revealed a fundamental divergence of point of view among mathematicians. There are those who take the "formalistic" point of view which maintains intact the Cantor theory. The opposing view, that of "intuitionism," not only declares many of the concepts of this theory as meaningless, but urges also a considerable modification of the logical bases of mathematics.

No purpose is served by mentioning a few of the many who have contributed to the development of the "Mengenlehre" during the last 50 years. Of the very extensive literature only a few books will be mentioned which do not presuppose a great deal of technical knowledge on the part of the reader. Among the best of these is *Einleitung in die Mengenlehre* by A. Fraenkel; the same author has written a short book entitled *Georg Cantor,* as a biographical introduction to the collected works of the pioneer in this field, now in process of publication. There is a large number of books in English

dealing in whole or in part with the "Mengenlehre," under such titles as *Theory of Sets*, *Theory of Sets of Points*, *Theory of Aggregates*. All books on the "Theory of Functions of a Real Variable" have chapters devoted to the subject. Most suitable for our needs is the book by Jourdain, *Contributions to the Founding of the Theory of Transfinite Numbers*. A proof of the equivalence theorem is found in most of the books alluded to; see, for example, Hobson, *Theory of Functions of a Real Variable*, Vol. I, 1907, p. 156; Borel, *Leçons sur la Théorie des Fonctions*, 1898, p. 104; and the first-mentioned book by Fraenkel, p. 54.

13. Coördinates in the plane — and a surprise. In spite of the difficulties involved in the study of the "Mengenlehre," several of

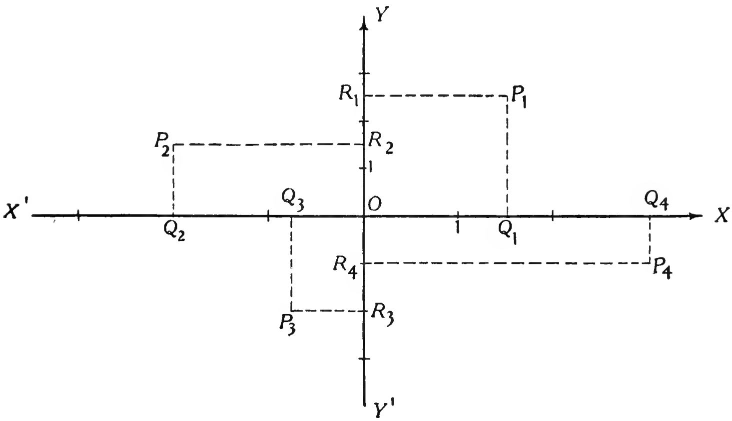


FIG. 3

its striking results are within our reach. We shall be able to penetrate far enough into Cantor's theory to understand why it did violence to "common sense" and to appreciate the refinement and enlargement of common notions which the study of infinite sets has made possible.

Let us recall first of all what is meant by the "coördinates" of a point in the plane. For the simplest form of coördinates we take two mutually perpendicular lines called the X-axis and the Y-axis. The point of intersection of the axes, O, is called the *origin of coördinates*. On each axis we designate one direction as the positive direction, and we adopt a unit of measurement (see Fig. 3). From an arbitrary point in the plane, P_1 (or P_2, or P_3, or P_4), we drop now

perpendiculars to each of the axes. Let the foot of the perpendicular on the X-axis be Q_1 (or Q_2, or Q_3, or Q_4); and let the perpendicular on the Y-axis land at R_1 (or R_2, or R_3, or R_4). Then the distance OQ_1, (OQ_2, OQ_3, OQ_4) measured in accordance with the unit and direction specified for the X-axis is the *x-coördinate* of P_1, (P_2, P_3, P_4), and the distance OR_1, (OR_2, OR_3, OR_4) measured in accordance with the unit and direction specified for the Y-axis is the *y-coördinate* of P_1, (P_2, P_3, P_4). The position of a point is now designated by writing its x-coördinate followed by a comma and that by its y-coördinate inside parentheses. Thus, with proper allowance for approximations in measurement, the point P_1 is designated as $(+1\frac{1}{2}, +2\frac{1}{2})$; the points P_2, P_3 and P_4 would receive the designations $(-2, +1\frac{1}{2})$, $(-\frac{3}{4}, -2)$, and $(+3, -1)$ respectively. The four parts into which the plane is divided by the X- and Y-axes are called *quadrants;* they are distinguished as the 1st, 2nd, 3rd and 4th quadrant, as indicated in Fig. 3. Points whose coördinates are given will lie in one of the 4 quadrants, according as these coördinates are positive or negative; the following table summarizes the facts:

y \ x	positive	zero	negative
positive	I	upper half of Y-axis	II
zero	right half of X-axis	origin	left half of X-axis
negative	IV	lower half of Y-axis	III

Thus, to every point P in the plane, corresponds a single ordered pair of real numbers (a, b) (see p. 4); and conversely. In other words, presumably now familiar, the set of points in the plane is equivalent to the set of ordered pairs of real numbers.

We consider now in particular the points in the plane both of whose coördinates are positive integers, or the equivalent set of ordered pairs of positive integers. Concerning them we shall prove the following theorem:

Theorem II. The set of points in the plane with positive integral coördinates is denumerable.

Let us designate the set of points in the plane with positive integral coördinates by P, the set of natural numbers by E. To prove our theorem we have merely to assign to each point of P a natural number in such a way that two different points will not have the same number. For in that case we surely have established a 1–1 correspondence between P and a part of E; but, on the other hand, it is clear that E is equivalent to a part of P, e.g. to the set of points in P whose x-coördinate is fixed at 1, i.e. to the set of points (1, 1), (1, 2), (1, 3), ... The equivalence theorem enables us then to conclude that the sets P and E are equivalent. To assign a natural number to every point of P we use a process which frequently finds application in different parts of mathematics. The

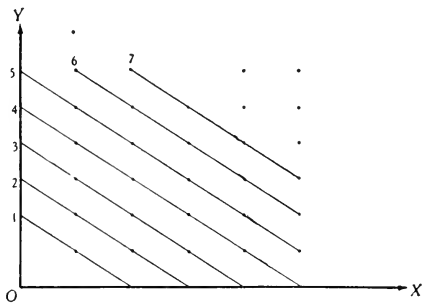


FIG. 4

dots in Fig. 4 suggest some of the points of P. The numbered lines connect all points for which the sum of the x- and y-coördinates is the same. On line *1*, $x + y = 2$; it contains *one* point of P, viz. (1, 1). On line *2*, $x + y = 3$; it contains *2* points of P, viz. (1, 2), (2, 1). On line *3*, $x + y = 4$; it contains *3* points of P, viz. (1, 3), (2, 2), (3, 1). Similarly, on each line there is a limited number of points of P; and these can be arranged in the order of magnitude of their x-coördinates. Thus, speaking generally, on line k, $x + y = k + 1$; it contains k points of P, viz. $(1, k)$, $(2, k - 1)$, $(3, k - 2)$ $\cdots$ $(k - 1, 2)$, $(k, 1)$. Moreover, one of these lines will pass through every point of P. Let us take a particular point of P, e.g. the point $A(9, 5)$; since the sum of the coördinates is 14, the point is on line 13. On the 12 lines which preceded, there are $1 + 2 + 3 + \cdots + 10 + 11 + 12 = 78$ points; since $A(9, 5)$ is the 9th point on line 13, it is the 87th point in the entire set P — we assign

to it the number 87. Consider now the perfectly general point $B(p, q)$ in which p and q are positive integers. Since the sum of the x- and y-coördinates is $p + q$, this point lies on line $p + q - 1$, and it is the pth point on this line; on the $p + q - 2$ lines which precede, there are $1 + 2 + 3 + \cdots + (p + q - 2)$ points. Therefore we assign to the point $B(p, q)$ the number $[1 + 2 + 3 + \cdots + (p + q - 2)] + p$.[1] For the numerical example just considered, $p = 9$, $q = 5$; the number assigned to the point $A(9, 5)$ was therefore $[1 + 2 + 3 + \cdots + 12] + 9 = 87$.

Thus we have assigned a single natural number to every point of P; moreover two different points $A(p, q)$ and $A_1(p_1, q_1)$ will not have the same number assigned to them. For,

(a) if $p + q = p_1 + q_1$, the points lie on the same line in Fig. 4. Hence if p is greater (less) than p_1, the number assigned to A will be greater (less) than that assigned to A_1; and if $p = p_1$, then $q = q_1$, so that A and A_1 coincide.

(b) If $p + q > p_1 + q_1$, A lies on a line which follows the line on which A_1 lies, so that its number surely exceeds that of A_1.

(c) If $p + q < p_1 + q_1$, the number given to A is less than that assigned to A_1. This completes the proof of the denumerability of the set of points in the plane with positive integral coördinates.

You, reader, who have followed it through step by step can not conceivably doubt its validity. Nevertheless you may feel not very much convinced. You may be willing to accept every step in the reasoning; but at the end of it, you may be quite at a loss, even though accepting the conclusion. This would not be surprising in view of the dominating control of your "common sense notions," grown out of your long experience with finite sets. But you are now dealing with infinite sets; a different "common sense" has to be built up, capable of giving guidance in experiences in which infinite sets are involved. This building up of an enlarged common sense is a process of growth; it may have to struggle for room. Hence the feeling of distress — growing pains.

[1] The reader may recognize that the numbers in the square brackets form an arithmetic progression and that their sum is equal to $\frac{(p + q - 2)(p + q - 1)}{2}$. If so, he will know that the natural number we have assigned to the point $P(p, q)$ is equal to $\frac{(p + q - 2)(p + q - 1)}{2} + p$; for the special case considered above $p = 9$, $q = 5$, and $\frac{(p + q - 2)(p + q - 1)}{2} + p = \frac{12 \cdot 13}{2} + 9 = 87$.

Another form in which the content of Theorem II may be put is the following: *The set of ordered pairs of natural numbers is equivalent to the set of the natural numbers.*

It is a corollary of this theorem that the set of points P, i.e. the set of points in the *plane* with positive integral coördinates is equivalent to the set of points on the X-axis with positive integral x-coördinates. Thus the theorem in a certain sense removes the distinction, in so far as points with positive integral coördinates are concerned, between the plane and the line, i.e. between a 2-dimensional and a 1-dimensional space. It is this which gives it a revolutionary tinge. The difficulties which stood in the way of acceptance of Cantor's theory, arose to some extent from the fact that it involved theorems of this sort which, even without restriction to points with positive integral coördinates, apparently abolish one of the distinctions between spaces of different dimensionalities. The next section will bring out this aspect still further.

14. Spaces of three, and more, dimensions. There are various ways in which Theorem II leads to other interesting results. First we shall consider an extension to 3 dimensions.

(a) The coördinates of a point in 3-dimensional space are determined most simply by means of 3 mutually perpendicular

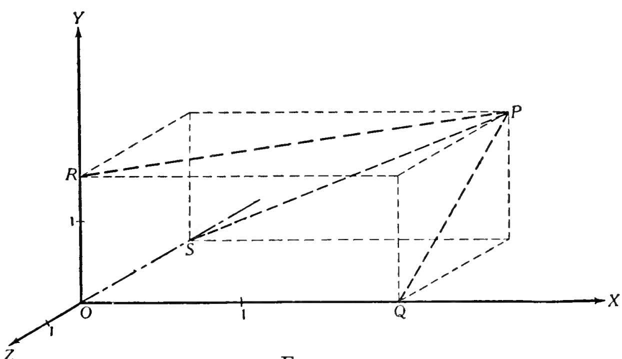


FIG. 5

planes called the coördinate planes, intersecting in a point O, the origin of coördinates; and, two by two, in lines called the coördinate axes, on which units and a positive direction are marked off. A point P in space being given, we drop perpendiculars from P on the coördinate axes, meeting these axes in Q, R and S, (see Fig. 5).

The coördinates of the point are then the numbers obtained when the lines OQ, OR and OS are measured in terms of the units and directions specified on the X-, Y- and Z-axis respectively. In the diagram of Fig. 5 the point P would be designated (allow for approximate character of result) by $(2, 1\frac{1}{4}, -1\frac{1}{4})$.

It can now be proved that the set of points in 3-dimensional space with positive integral coördinates is denumerable; in other words, we assert

Theorem III. The set of ordered triads of natural numbers is denumerable.

The essential step in the proof of this proposition is as follows: The 1–1 correspondence between the set of ordered pairs (p, q) and the set of numbers

$$n = \frac{(p+q-1)(p+q-2)}{2} + p,$$

which was established in the proof of Theorem II carries with it a 1–1 correspondence between the set of ordered triads (p, q, r) and the set of ordered pairs

$$\left[\frac{(p+q-1)(p+q-2)}{2} + p, r\right],$$

i.e. between the set of ordered triads (p, q, r) and a part of the set of ordered pairs (n, r). The reader should have no difficulty in completing the proof from this point on (see *17*, 9, p. 30).

Moreover, he should be able to obtain the further result that the sets of ordered tetrads of positive integers, of ordered pentads of positive integers, and so forth are all denumerable. By analogy with the geometrical form in which the contents of Theorems II and III can be stated, it is customary to cast these additional results in geometrical form also, as follows: The sets of points in 4-dimensional, 5-dimensional space, etc. with positive integral coördinates are denumerable, i.e. equivalent to the set of such points on a line (see *17*, 10, p. 31). If extended so as to remove the restriction to positive integral coördinates, as was done by Cantor, these theorems establish a property common to spaces of different dimensionalities.[1]

[1] It should be remembered that "point in 4-dimensional (5-dimensional) space" is merely a geometrical form of saying "tetrad (pentad) of real numbers."

(b) Any infinite set which is a part of the set of ordered pairs of natural numbers is denumerable. An important example of this is found in the set of proper fractions reduced to lowest terms. Recalling Definition I and footnote 1 on page 18, we know that a proper fraction reduced to lowest terms is a number of the form $\frac{p}{q}$, in which p and q are natural numbers, without common factors, $p < q$; it is a rational number, less than 1 (see p. 4). There is therefore an evident 1–1 correspondence between the proper fractions and the ordered pairs of relatively prime natural numbers, (p, q), in which $p < q$. Consequently *the set of proper fractions is denumerable.* The details of the proof are left to the reader (see *17*, 11, p. 31).

15. Inequality — its symbols and its rules. The last result gives us new insight into the possibilities of a denumerable set. The 1-dimensional denumerable sets with which we have dealt so far have been composed of points with a fixed distance between any two successive points, or of numbers of which the mutual differences did not fall below a certain fixed amount. In contrast to these sets, the set of proper fractions has the property that between any two of its elements there exists another one; i.e. between any two proper fractions we can insert another one. For, if $\frac{p}{q}$ and $\frac{p_1}{q_1}$ are proper fractions, then $\frac{1}{2}\left(\frac{p}{q} + \frac{p_1}{q_1}\right)$ is also a proper fraction; and moreover, if $\frac{p}{q} < \frac{p_1}{q_1}$, then $\frac{p}{q} < \frac{1}{2}\left(\frac{p}{q} + \frac{p_1}{q_1}\right) < \frac{p_1}{q_1}$, i.e. the new fraction lies between the two given ones. The proofs of these statements are very simple, provided we know the fundamental laws governing operations on inequalities.

Here is another sidepath which invites us. It will be worth our while to follow it for a short distance, partly because it will enable us to prove the above assertion about common fractions (which the reasonable reader will surely accept as reasonable), partly because it affords a valuable contrast with things already known, and partly because it will give us a valuable tool for future use.

In school mathematics inequalities are rarely encountered, except in a few propositions in geometry. But it can readily be understood why inequalities should play at least as important a part

as equalities in the applications of mathematics; for both in nature and in human society, inequality is the rule, equality the exception, only infrequently realized. What then are the fundamental rules concerning inequalities? To such a direct question fits only a direct answer. The fundamental rules governing inequalities are the following:[1]

1. If $a \leqq b$, and $b \leqq c$, then $a \leqq c$.
2. If $a < b$, and $a_1 \leqq b_1$, then $a + a_1 < b + b_1$.
3. If $0 < a < b$, and $0 < a_1 \leqq b_1$, then $aa_1 < bb_1$, and $\dfrac{a}{b_1} < \dfrac{b}{a_1}$.

It would take a good many words if these rules were to be written out without the use of symbols. But words are *spoken* rapidly enough; the reader is therefore urged to state these rules fully in words, and to reflect a bit on their meaning. Does anyone ask for the authority for these rules, by whom they were imposed, their domain of validity, etc.? He must remain unanswered at present and hold his question in reserve for a while (see Chapter X).

Armed with these laws, we now return to the set of proper fractions:

(a) From $0 < p < q$, follows $pq_1 < qq_1$, by (3); and
from $0 < p_1 < q_1$, follows $p_1q < qq_1$, by (3).

Hence, from the two hypotheses follows $pq_1 + p_1q < 2qq_1$, by (2). Therefore, if $\dfrac{p}{q}$ and $\dfrac{p_1}{q_1}$ are proper fractions, then

$$\frac{pq_1 + p_1q}{2qq_1} = \frac{1}{2}\left(\frac{pq_1}{qq_1} + \frac{p_1q}{qq_1}\right) = \frac{1}{2}\left(\frac{p}{q} + \frac{p_1}{q_1}\right)$$

is also a proper fraction.

(b) From $\dfrac{p}{q} < \dfrac{p_1}{q_1}$, follows $\dfrac{2p}{q} = \dfrac{p}{q} + \dfrac{p}{q} < \dfrac{p}{q} + \dfrac{p_1}{q_1}$, by (2);

and also $\qquad \dfrac{p}{q} + \dfrac{p_1}{q_1} < \dfrac{p_1}{q_1} + \dfrac{p_1}{q_1} = \dfrac{2p_1}{q_1}$, by (2).

[1] The meaning of the symbols $<$ and $>$ has already been recalled (see p. 9). It happens often that combination symbols like $\leqq$ or $\geqq$ are used; they stand respectively for "less than or equal to" and "greater than or equal to." Very useful is the continued inequality which appears in forms like the following: $a < b < c$; it is an abbreviated form for "$a < b$ and $b < c$" and has the advantage of placing the symbol b between the symbols a and c, when it represents a number between a and c in magnitude.

This shows that $\frac{p}{q} < \frac{p_1}{q_1}$ implies the relation

$$\frac{p}{q} < \frac{1}{2}\left(\frac{p}{q} + \frac{p_1}{q_1}\right) < \frac{p_1}{q_1}, \text{ by (3).}$$

Thus we have established the property of the set of proper fractions referred to above. It is a noteworthy fact that in spite of possessing this property, the set is denumerable.

16. A non-denumerable set. These results suggest once more the question first asked on page 18: "Are there any non-denumerable sets?" We shall now justify the affirmative answer which we gave at that point, although not by means of the geometrical instance there mentioned. Instead we shall prove:

Theorem IV. The set of all proper decimal fractions is not denumerable.

Proof. A proper decimal fraction is a decimal fraction whose value is between 0 and 1; it may be terminating or non-terminating. If we agree to write *zeros* after the last decimal in a terminating decimal fraction, we can consider all the fractions under discussion as non-terminating. The numbers of our set are therefore of the form 0.57803259462 . . . in which the decimals continue without end although all of them from a certain point on may be zeros. The theorem was one of the early results obtained by Cantor in his development of the theory of sets. It remains to notice how simply he proved it. For, says Cantor, suppose that the set of all these decimal fractions were denumerable.[1] We could then think of them as arranged in order, a first one, a second one, a third one, and so forth. Let us suppose that the successive places of the first one were filled by digits $a_{11}a_{12}a_{13}a_{14}a_{15}$ . . . each of the letters a_{11}, a_{12}, etc. standing for some one of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. Let us suppose that the successive places of the second number on the list were filled by the digits $a_{21}a_{22}a_{23}$ etc., those of the third by $a_{31}a_{32}a_{33}a_{34}$ etc., and so forth; every proper decimal fraction would have to occur somewhere in the list. Cantor proved that this is impossible by showing that there is *at least one decimal fraction*

[1] A terminating decimal can be represented in two ways as a non-terminating decimal; for example, .14000 . . . represents the same number as .13999 . . . . We suppose that the set is extended so as to include both representations of each terminating decimal. If the set of all proper decimals were denumerable, this extended set would also be denumerable (compare *10* and *13*).

which can not occur in the list. Such a fraction he constructed as follows: Let A_1 be a digit which is *different* from a_{11}; let A_2 be a digit different from a_{22}, A_3 different from a_{33}, A_4 different from a_{44}, etc. Consider now the decimal fraction whose successive places are filled by $A_1, A_2, A_3, \ldots$; it can not occur in the list. For it surely differs from the first one in the list in its first place, from the second one in its second place, from the third one in its third place, and so forth. Thus the theorem is proved.

The set of all proper decimal fractions is frequently called the "*continuum* between 0 and 1." In the continuum we have thus learned to know an infinite set not equivalent to the set of natural numbers. And now comes the question whether every infinite set is equivalent to one or the other of these two sets? Also whether it is possible to set up relations of order between non-equivalent infinite sets. These and many other questions belong to the general theory of sets; they are dealt with in the books mentioned in *12*. They have no place in our introductory treatment. With the proof of Theorem IV we have to close this part of our journey; a few exercises are desirable however to fix some of the concepts.

17. To secure the treasures.

1. Indicate on a sheet of coördinate paper the position of the points determined by the following pairs of coördinates: $A(-3, 3)$; $B(2, 1)$; $C(5, 3)$; $D(4, -4)$; $E(-3, -4)$; $F(-1, -3)$; $G(0, 0)$; $H(0, -1)$; $I(2, 0)$.

2. Determine a relation which holds
   (a) between the x- and y-coördinates of A, D, G.
   (b) between the x- and y-coördinates of B, E, H.
   (c) between the x- and y-coördinates of C, F, I.

3. Locate three or four points whose coördinates satisfy the equation $x + y = 5$; also for the equations $2x - y = 7$ and $x + 3y = -2$.

4. Establish a 1–1 correspondence between the points with positive integral coördinates on the line $2x - y = 7$, and the natural numbers.

5. Prove Theorem II by the method suggested by the diagram in Fig. 6. Do other methods occur to you by which this theorem might be proved?

6. Show that, if $\frac{p}{q}$ and $\frac{p_1}{q_1}$ are proper fractions, then $\frac{3pq_1 + 5p_1q}{8qq_1}$ is a proper fraction which lies between them.

7. Prove that the set of all Pythagorean triples is denumerable.

8. Show that if $a < b < 0$, and $a_1 < b_1 < 0$, then $aa_1 > bb_1 > 0$.

9. Complete the proof of Theorem III, p. 26.

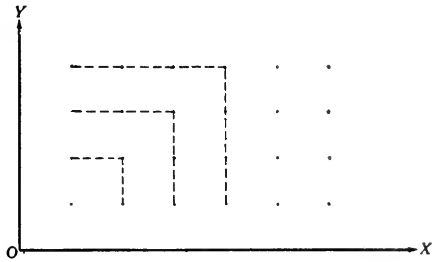


FIG. 6

10. Prove that the set of all ordered tetrads of positive integers is denumerable.

11. Complete the proof that the set of all proper fractions is denumerable (see p. 27).

12. Prove that the set of all decimal fractions, terminating and nonterminating, in which only the digits 0, 1, 2, 3, 4 occur, is not denumerable.

13. Prove that an infinite subset of a denumerable set is itself denumerable.

14. Prove that a subset of a non-denumerable set may be finite, denumerable, or non-denumerable.

CHAPTER III

THE SECOND EXCURSION — WALLED CITIES

The train of regulated thoughts is of two kinds; one, when of an effect imagined we seek the causes or means that produce it; and this is common to man and beast. The other is when imagining anything whatsoever we seek all the possible effects that can by it be produced, that is to say, we imagine what we can do with it when we have it. Of which I have not at any time seen any sign but in man only; for this is a curiosity hardly incident to the nature of any living creature that has no other passion but sensual, such as are hunger, thirst, lust and anger. — Thomas Hobbes, *Of Man*.

It will be recalled that in the proof of Lemma 1, on page 7, we concluded, as any child would, that since $q = f\sqrt{r_1^2 - p_1^2}$, q has the factor f. Let us, now we are growing to man's stature, put childish things away and examine this argument a little more closely. Would we say that 5 has the factor 7 because $5 = 7 \times \frac{5}{7}$? Would we call 3 a perfect square on the basis of the equation $3 = \sqrt{3} \times \sqrt{3}$? What do we really mean by the factors of a number? Students in high schools are set to factoring in their freshman year. It is a favorite pastime for them. But what does it all mean? A clarification of our views on this point is called for — the present chapter has this for its purpose; at a later time (see Chapter VII) we shall return to the question whether the argument in the proof of Lemma 1 is valid.

18. Reminiscence of childhood. It should be clear that if the identity $5 = 7 \times \frac{5}{7}$ were used to justify the statement that 7 is a factor of 5, then every integer would be a factor of every other one; and if the fact that $3 = \sqrt{3} \times \sqrt{3}$ were to be accepted as basis for the statement that 3 is a perfect square, then every integer would be a perfect square. The problem of factorization would be completely solved; at the same time it would lose all interest and become unworthy of the serious attention of students or teachers in high schools. (What a relief that would be!)

But we are not in this state. It has been tacitly understood that to factor a given natural number a means to find two or more natural numbers whose product is equal to a; i.e. as "factors" of a

natural number we admit only natural numbers. When we say that the integer b is a perfect square, we mean that there exists an *integer* c such that $b = c \times c$. These simple observations reveal the fact that underlying the operation of factoring an integer there has been assumed a perfectly definite set of numbers, viz. the set of integers, within which the factors are to be found. It is conceivable that a different set of numbers might be agreed upon for the factors. As soon as the set of numbers has been specified to which the factors are to be restricted, the problem of factoring becomes a definite one, although not always interesting and significant; without such specification the problem is meaningless.

The set of numbers to which the factors of numbers of a certain type are to belong, may be chosen arbitrarily, except for the condition that it must be possible to multiply together any two numbers of this set. For convenience of terminology we shall speak of the "product set" and the "factor set" in referring respectively to a given set of numbers and to the set of numbers admitted as their factors. Let us now consider a few examples.

(a) If the "product set" is the set of positive integers and the "factor set" is the set of positive and negative integers, the factorization problem is definite and significant.

(b) If the "product set" is the set of positive integers and the "factor set" the set of natural numbers, the problem is definite, not meaningless, but without interest. For factoring is never possible in this case, inasmuch as the product of two natural numbers is a natural number, not a positive number.

(c) If the "product set" consists of the negative integers and zero and the "factor set" of the negative and positive integers and zero, the problem is definite and significant.

The definitions of prime numbers and of pairs of relatively prime numbers given on pages 7 and 8, tacitly assume that the "factor set" is the set of natural numbers. When a new factor set is introduced, the definition of prime factor will in general have to be modified.

The purposes for which the factors of integers are needed are best served when the factoring process is carried out completely, i.e. until a set of prime factors has been reached. The corresponding problem with respect to other factor sets acquires meaning when the prime factors have been specified.

19. What does "impossible" mean? Let us now look at some other types of factoring familiar from school days. One of the

things everybody remembers is that $a^2 - b^2 = (a + b)(a - b)$, but that $a^2 + b^2$ can not be factored. No wide-awake young person should accept a statement like the last; for clearly

$$a^2 + b^2 = \frac{(a + b) \times (a^2 + b^2)}{a + b},$$

or, if preferred, $a^2 + b^2 = (a + bi)(a - bi)$.[1] What then is meant by the statement that $a^2 - b^2$ can, but $a^2 + b^2$ can not be factored? The answer to this question is to be found in the discovery of a "factor set," i.e. a set of admissible factors. With reference to a definite factor set the meaning of the statement "$a^2 - b^2$ is factorable, but $a^2 + b^2$ is not" is perfectly clear — whether it is true or false is quite another matter.

Factorization in high school deals largely with "polynomials in one or more variables with real coefficients." The factor set which is tacitly assumed is the set of such polynomials; and a prime factor is a polynomial of this kind whose only factors are the polynomial itself together with one or more real numbers. Within this set, the factors $a + b$ and $a - b$ are admissible, but $a + ib$ and $a - ib$ are not admissible because i is not a real number. The factorization $3a^2 - 2b^2 = (\sqrt{3}a + \sqrt{2}b)(\sqrt{3}a - \sqrt{2}b)$ is certainly admissible in this set; it would not be admissible if the factor set were the set of polynomials with rational coefficients. To resolve a polynomial into prime factors is a definite undertaking only when the factor set and the prime factors have been designated. Even after this has been done, it is not always easy to decide whether or not a given polynomial is prime, or to find its factors in case it is not prime; until this has been done, the problem is without meaning.

20. Old questions in a new light.[2]

1. Factor the following numbers into prime factors, the factor set and the prime factors being specified for each group:

(a) $+ 372$, $+ 532$, $- 441$ (Positive integers; a prime factor is a positive number, $\neq 1$, whose only factors are itself and $+ 1$.)

(b) $+ 364$, $+ 795$, $- 627$ (Negative integers; a prime factor is the negative of a prime factor in (a)).

[1] The reader will remember that i is a number with the property $i^2 = -1$.

[2] The reader will experience some peculiar difficulties in a few of these problems; they arise from the fact that the product sets sometimes contain numbers which can not be obtained by the multiplication of elements of the factor set (see 9).

(c) $+\,819$, $+\,2145$, $-\,1309$ (Positive and negative integers; a prime factor is an integer whose only factors are itself or its negative and $+\,1$ or $-\,1$).

2. Discuss the factorization of negative integers for each of the three factor sets designated in problem 1.

3. Show that the factorization of rational numbers is without interest when the factor set consists of the set of rational numbers.

4. Determine the prime factors of the following numbers, the factor set and prime factors being indicated for each group:

(a) 150, 490, 726, 1155 (2, 3, 5, 7, 11; prime factors as in Def. II);

(b) 910, 1365, 462, 546 (2, 3, 5, 7, 13; prime factors as in Def. II).

5. Prove that the complete factorization of elements of any product set with respect to the set of integers from 1 to 10 is the same as with respect to the set 2, 3, 5, 7.

6. Prove that the complete factorization of positive integers with respect to the set of positive integers is the same as with respect to the set of prime positive integers.

7. Factor each of the following numbers completely, the factor set and the prime factors being indicated for each group:

(a) 28, 220, 1810, 2464 ($4, 7, \cdots, 1 + 3n, \cdots$; a number is prime when its only factors in the set are itself and 1);

(b) 65, 105, 693, 4389 ($5, 9, \cdots, 1 + 4n, \cdots$; primes as in (a)).

8. Factor the following polynomials, the factor set and the prime factors being as indicated:

(a) $8a^3 - 27b^3$, $8a^3 + 27b^3$, $6a^2 + 13ab + 4b^2$ (Polynomials with integral coefficients; prime factors as in *19*);

(b) $4a^4 - 9b^4$, $4a^2 - 5b^2$, $4a^6 - 9b^6$ (as in (a));

(c) $4a^4 - 9b^4$, $8a^3 - 27b^3$, $4a^2 - 5b^2$ (Polynomials with real coefficients; prime factors as in (a)).

9. Discuss the difficulties which arise in factorization when the product set contains numbers which can not be obtained by the multiplication of elements of the factor set (see footnote 2 on p. 34).

10. Factor the following polynomials when the factor set is the set of polynomials with real, or normal coefficients (a "normal" number is a number of the form pi, where p is a real number and $i^2 = -1$), and the prime factors are specified as in *19:*

(a) $4a^2 + 9b^2$, (b) $9a^2 - 3c^2$, (c) $5a^2 + 2b^2$, (d) $4a^4 - 9b^4$, (e) $a^2 + ab + b^2$.

11. Factor the following polynomials when the factor set is the set of polynomials of *even* degree with real, or normal coefficients, and the prime factors are as specified in *19:*

(a) $4a^4 - 9b^4$, (b) $4a^4 + 9b^4$, (c) $a^3 - b^3$.

12. Factor the following polynomials when the factor set is the set of

polynomials of even degree with integral coefficients, the prime factors being again as in *19:*

(a) $a^4 - b^4$, (b) $a^6 + b^6$, (c) $a^6 - b^6$.

21. Rational operations. The preceding discussion and exercises reveal the importance for factorization of considering collections or sets of numbers, rather than individual numbers. It is one of the characteristics of the modern development in the theory of numbers that it leads to the study of *collections of numbers* with reference to the various operations which arise. We must now consider some of the familiar operations from this point of view. We have already encountered different collections of numbers, e.g. rational numbers, normal numbers, real numbers, etc. It is desirable that we should now learn to understand clearly the meaning of the words we have used to designate sets of numbers, and which have been used rather loosely thus far. In considering them, we shall at the same time be concerned with some of their properties in regard to the operations, *addition, subtraction, multiplication,* and *division.* These four operations are usually referred to as the *rational operations;* but, for reasons which will become clear presently, (see *23*), *division by zero* is excluded from the rational operations.

22. Adventure in arithmetic. At the beginning was the whole number. The set of "natural numbers," 1, 2, 3, . . . lies at the basis of all mathematical analysis. Excessive and mystical veneration of these entities has not been infrequent in the past, nor is it absent at the present time. A great deal of interesting material as to the origin of the concept of number and of the symbols used to designate them can be found in the books by Dantzig and by Conant, mentioned on page 14, (see also J. Pierpont, *Theory of Functions of a Real Variable,* Vol. I, pp. 1–5; for a fundamental treatment of the natural numbers the reader is referred to *Theoretische Arithmetik,* by O. Stolz and J. A. Gmeiner). To the German mathematician Leopold Kronecker (1823–1891) is attributed the saying: "the whole number was created by God, everything else is man's handiwork" which gives expression to a belief in the fundamental character of the natural numbers.[1] We make now the following observations concerning them:

[1] For some interesting aspects of the veneration of numbers, see E. T. Bell, *Numerology* (Baltimore, 1933).

1. If any two natural numbers a and b are given there exists just one third natural number c which is equal to their *sum:* $c = a + b$.

2. If any two natural numbers a and b are given there exists just one third natural number d equal to their *product:* $d = ab$.

In other words, the operations *addition* and *multiplication* can be carried out uniquely within the set of natural numbers, *without any restriction.* This is sometimes expressed in the convenient phrase: The set of natural numbers is *closed* with respect to addition and multiplication.

To add the numbers a and b means to *find* the number c, such that $a + b = c$; to multiply a and b means to *find* the number d, such that $ab = d$. Many hours are spent by every "civilized" human being in attempts to learn the process by which these numbers c and d are to be located. Let no one be disturbed at this point by the fear that the old horror of multiplication tables is to descend upon him once more. We are not now concerned at all with the process; in spite of all popular misconceptions, mathematicians are not walking tables of multiplication. As a matter of fact they are not very often dealing with numbers as individuals. What we *are* concerned with is the assertion that (a) no matter what natural numbers a and b may be given, their sum and product always exist within the set of natural numbers: and (b) there is *only one* sum and *only one* product for every pair of natural numbers. This fact, that problems in addition and multiplication of natural numbers always have *one* and *only* one answer has been the basis of much argument both *for* and *against* the teaching of arithmetic. The writer does not pretend to any qualification in matters of pedagogy. But it has always seemed to him that not half enough advantage has been drawn from the element of adventure which this uniqueness of the result lends to the search for sums and products.

In sharp contrast with this state of affairs is the condition which obtains when we attempt to invert the operation of addition within the set of natural numbers, i.e. if, when a and c are given natural numbers, we seek to determine the natural number b so that $a + b = c$, or, in familiar words if we want to subtract a from c. Here condition (a) is not fulfilled; only if $a < c$ can such a natural number b be found. On the other hand, *if* $a < c$, then condition (b) does hold. If and only if $a < c$, can the subtraction of a from c

be carried out in the set of natural numbers; but never can there be more than *one* difference between two natural numbers.

A similar situation holds with respect to the operation division, which is the inverse of multiplication. To divide d by a means to find the natural number b so that $a \times b = d$. This is not possible unless d is a multiple of a, i.e. unless d occurs among the numbers $a, 2a, 3a, 4a, \ldots$; if it does, there is only one such number b.

23. The rational numbers. A fundamental principle guiding the development of mathematical concepts is the extension of sets of elements in such a way as to make the operations to be performed on these elements possible without exception. This principle grows in part out of the need for such extensions experienced in the application of the mathematical concepts to various concrete experiences.

The natural numbers suffice for the needs of counting objects in any finite collection. But for measurement they are not sufficient. To weigh a quantity of wheat, to determine the length of a bar of metal, to determine temperatures, heights of mountains or depths of waters, the natural numbers are inadequate.[1] No matter what unit of weight is used, it is not to be expected that the weight of all objects that are to be weighed can be expressed in terms of this unit by a whole number; and no one unit of length can be used for all length measurements. Moreover no matter what temperature is used as a starting point, the natural numbers are not likely to be sufficient to express all temperatures. A similar lack of adequacy is found in many processes of measuring, when only the natural numbers are at our disposal.

It would be an interesting and valuable exercise to show that these difficulties encountered in measuring arise from the fact that in the set of natural numbers the inverse operations, subtraction and division, are not always possible; in another terminology, that the set of natural numbers, although closed under addition and multiplication is not closed under subtraction and division. Leaving this exercise to the reader, we shall now consider a new set of numbers, in which not only addition and multiplication, but also subtraction and division, i.e. all the rational operations are possible without restriction, or rather with but one single exception. This set of numbers, called the *set of rational numbers*, consists of the

[1] A valuable discussion of these matters is found in *Counting and Measuring*, by H. von Helmholtz, translated by Charlotte Lowe Bryan (New York, 1930).

positive and negative integers, the positive and negative fractions, and zero.

The critical reader will doubtless want to know where these numbers come from. The answer may be found in various places; but to reach any of these a considerable journey is required, which it would not be sensible for us to undertake now — we have another objective before us. The reader is referred to J. Pierpont, *Theory of Functions of a Real Variable*, Vol. I, pp. 1–19, to E. Landau, *Grundlagen der Analysis*, pp. 1–42, to O. Perron, *Irrationalzahlen*, pp. 1–5, or to B. Russell, *Introduction to Mathematical Philosophy*, pp. 1–20, 63–77.[1]

We shall proceed, on the assumption that the rational processes as applied to rational numbers are familiar from school days, to indicate some of the properties of this set of numbers:

1. Every rational number is the quotient of two integers; these integers may be positive or negative. For the positive and negative fractions this property is obvious; it becomes also obvious for the integers, as soon as we notice that the integer a is equal to the fraction $\frac{a}{1}$.

Conversely, the quotient of two integers is a rational number. For, if the divisor be 1, the quotient is a positive or negative integer and hence a rational number; and if it is an integer different from 1, the quotient is a common fraction, again a rational number. But we have to note an exception! Suppose that the divisor were 0. To divide a by 0 means, according to *22*, to find a rational number b such that $0 \times b = a$. Now if $a \neq 0$, no such rational number b can be found, since $0 \times b = 0$, no matter what rational number b may be; and if $a = 0$, then every number will do. In neither case do we obtain a satisfactory result. For this reason, *division by zero is excluded from the rational operations* (see *22*).

2. The sum and difference, the product and quotient of two rational numbers is a rational number.

Let the two rational numbers be $\frac{p_1}{q_1}$ and $\frac{p_2}{q_2}$, where p_1, q_1, p_2 and q_2 are integers, positive or negative, where p_1 and q_1 are relatively prime and also p_2 and q_2. Then

[1] Compare *39*, 3, p. 69.

$$\frac{p_1}{q_1} + \frac{p_2}{q_2} = \frac{p_1q_2 + p_2q_1}{q_1q_2}, \quad \frac{p_1}{q_1} - \frac{p_2}{q_2} = \frac{p_1q_2 - p_2q_1}{q_1q_2},$$

$$\frac{p_1}{q_1} \cdot \frac{p_2}{q_2} = \frac{p_1p_2}{q_1q_2}, \quad \frac{p_1}{p_2} \div \frac{q_1}{q_2} = \frac{p_1q_2}{p_2q_1}.$$

But $p_1q_2 + p_2q_1$, $p_1q_2 - p_2q_1$, q_1q_2, p_1p_2, p_1q_2 and p_2q_1 are all integers; hence we have shown that the *set of rational numbers is closed under all rational operations.*

24. Alice in Wonderland. The meaning of this concept of "closure" should be well understood. If a set of numbers is closed under certain operations, it is like a self-perpetuating corporation, or better, like a walled city whose gates are closed and guarded. There is no way out for him who obeys the prevailing laws; there is no way of getting out of the set of numbers so long as we limit ourselves to the operations with respect to which the set is closed. It is comfortable and conservative; no surprises in the form of unsolvable problems arise, but neither is there an inkling of new worlds beyond the walls.

Enlargement of experience frequently, though not always, awaits those who manage to find an opening in the city walls. Just so, when we discover an operation with respect to which a set of numbers is not closed, we have the possibility of enlargement. We seek then to embed the set in a new and extended set in such a way that the operation can be carried out without restrictions in this new set. This is what happens when the operations of subtraction and division lead us from the set of natural numbers to the set of rational numbers.

It is worthy of note that the extension is made through the inverses of operations with respect to which the set of natural numbers is closed. For indeed, inversion is a very fertile source of extensions. One does not fully understand the significance of an operation until one has considered its inverse. It is this which led the famous German mathematician, C. G. J. Jacobi (1804–1851) to the dictum: "Man muss immer umkehren" (We must always invert).[1]

There is significance in this principle far beyond the domain of mathematics. To appreciate the significance of an act it has to be

[1] We are using the word "invert" as a convenient translation of the German "umkehren," although the two words are by no means exact equivalents.

considered not merely from the point of view of the actor, but also from that of those whom the act is acted upon.

To foresee the effects of an act imaginatively is an important element in understanding it. If we understand the word "invert" somewhat freely, we can say that Lewis Carroll understood the significance of this principle (it is not for nothing that he was a mathematical tutor) when he chronicled the experiences of *Alice in Wonderland* and *Through the Looking Glass*. The importance of looking at the world from an unusual point of view can not be overestimated. There would be a great deal of value in the experience of a person who could stand on his head for a day at a street corner in one of our big cities. International relations would stand under a constellation quite different from the one that guides their destinies now if statesmen as well as ordinary citizens could understand the principle of inversion, if they would judge a proposed agreement between different peoples from the points of view of each of the parties involved. And, what is the Golden Rule but an admonition to apply the principle of inversion to the relations between human beings.

We have come a long way, from the study of rational numbers to the Golden Rule; but the connection was direct, the path without obstacles. We should traverse it backwards and forwards, several times, until we have established the connection firmly in our own minds. And then we can go on.

In the system of rational numbers the distinction between addition and multiplication on the one hand, subtraction and division on the other is no longer essential; the direct and the inverse have passed over into one another. For if b is a rational number, so are $-b$ and $\frac{1}{b}$. Consequently, since to subtract b from a is the same as adding $-b$ to a, every operation of subtracting is converted into an addition. Similarly since to divide a by b is equivalent to multiplying it by $\frac{1}{b}$, every division is converted into a multiplication.

We have therefore the complete content of property 2, stated on page 39, if we say:

Theorem V. The set of rational numbers is closed under addition and multiplication.

A large part of the time and energy of the nation's youth is spent in verifying over and over again the validity of Theorem V. For

when the long-suffering youngsters are asked by a textbook writer to reduce $\dfrac{\frac{3}{5}-\frac{4}{7}}{\frac{2}{3}+\frac{1}{5}}$, they are expected to carry out the different rational operations here indicated and to determine the rational number (i.e. the quotient of two integers) to which they lead. The troubles most people have with this sort of thing arise from a lack of understanding of what such a jumble of numbers and lines really means; if looked upon as a design, one can readily learn to unravel it. But these problems are not fully understood until it has become clear that they are, all of them, illustrations of *one fact*, viz. that the set of rational numbers is closed under rational operations.

25. Arithmetic in a nutshell. There are many other important properties of the set of rational numbers besides those discussed in *24*. We recall the fact established in *15* that between any two rational numbers another one can be inserted.

We shall put together now in systematic order some of these properties of rational numbers in conjunction with the operations addition and multiplication; they constitute the basis of all our arithmetic. Let us denote by R the set of rational numbers; then we can say "a belongs to R" instead of "a is a rational number."

1. If a and b belong to R, there exists a unique element of R equal to $a + b$.

2. If a and b belong to R, there exists a unique element R equal to ab.

These are our old friends the closure properties.

3. If a and b belong to R, then $a + b = b + a$.

4. If a and b belong to R, then $ab = ba$.

These two properties are known as the *Commutative law* of addition and multiplication respectively. They will strike the reader as foolishly simple and insignificant; but we must remember that children and fools have the reputation of speaking the truth — and that is all we are concerned with now.

5. If a, b and c belong to R, then $(a + b) + c = a + (b + c)$.

6. If a, b and c belong to R, then $(ab)c = a(bc)$.

These two properties are known as the *Associative law* of addition and multiplication respectively. The reader should note carefully the meaning of the use of the parentheses. He should also consider whether such a symbol as $(a + b) + c$ would have any meaning if the closure theorem did not hold. It follows from 5

and 6 that the parentheses can safely be omitted, because the symbols $a + b + c$ and abc are unambiguous in meaning.

7. If a, b and c belong to R, then $a(b + c) = ab + ac$.

This is the first and only one of the properties here listed in which the two operations addition and multiplication are combined. It is called the *Distributive law* of multiplication; it says that the multiplication of a sum by a number is distributed over the elements which make up this sum. In combination with 4, we obtain from 7, that $(a + b)c = ac + bc$.

8. There exists in R a single element z such that, no matter what element of R a is, it is always true that

$$a + z = a.$$

9. There exists in R a single element u such that, no matter what element of R a is, it is always true that

$$au = a.$$

Properties 8 and 9 will strike the reader as a sophisticated way of speaking about 0 and 1. But he may not have been aware of just what those familiar symbols meant in the operations on rational numbers, nor of the interesting fact that there is but a single rational number of each kind.

10. No matter what element of R a is, there always exists a single rational number n_a such that

$$a + n_a = z.$$

11. No matter what rational number a is, *provided that it be not* z, there always exists a single rational number r_a such that

$$ar_a = u.$$

It must be clearly understood that the elements n_a and r_a whose existence is asserted in 10 and 11 are not unique for the whole of R, as were z and u, but that for every a of R, there is one n_a and one r_a. The symbol n_a is used to suggest "negative of a"; and the notation r_a serves to suggest "reciprocal of a."

The last two properties merely bring out the fact that in the set of rational numbers the inverse operations subtraction and division are possible without restrictions, except for division by 0.

26. Fields. These properties of the set R of rational numbers will gain significance if we test them on other sets of numbers. Which of them hold if we use the set of natural numbers instead

of *R*? Clearly, 1, 2, 3, 4, 5, 6, 7 and 9 do hold; but 8, 10 and 11 do not.

For the set consisting of 0 and the positive integers from + 1 to + 10, very few of the properties hold if addition and multiplication are understood in the ordinary way. But if we agree to reduce by 10 any sum or product which exceeds 10 but not 20, by 20 any sum or product which exceeds 20 but not 30 and so forth, we would fare somewhat better with our properties. We would have the following addition table:

	0	1	2	3	4	5	6	7	8	9	10
0	0	1	2	3	4	5	6	7	8	9	10
1	1	2	3	4	5	6	7	8	9	10	1
2	2	3	4	5	6	7	8	9	10	1	2
3	3	4	5	6	7	8	9	10	1	2	3
4	4	5	6	7	8	9	10	1	2	3	4
5	5	6	7	8	9	10	1	2	3	4	5
6	6	7	8	9	10	1	2	3	4	5	6
7	7	8	9	10	1	2	3	4	5	6	7
8	8	9	10	1	2	3	4	5	6	7	8
9	9	10	1	2	3	4	5	6	7	8	9
10	10	1	2	3	4	5	6	7	8	9	10

and the multiplication table would run as follows:

	0	1	2	3	4	5	6	7	8	9	10
0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7	8	9	10
2	0	2	4	6	8	10	2	4	6	8	10
3	0	3	6	9	2	5	8	1	4	7	10
4	0	4	8	2	6	10	4	8	2	6	10
5	0	5	10	5	10	5	10	5	10	5	10
6	0	6	2	8	4	10	6	2	8	4	10
7	0	7	4	1	8	5	2	9	6	3	10
8	0	8	6	4	2	10	8	6	4	2	10
9	0	9	8	7	6	5	4	3	2	1	10
10	0	10	10	10	10	10	10	10	10	10	10

By the aid of these tables it is easy to verify that properties 1, 2, 3, 4, 5, 6, 7, 8 and 9 all hold, if 0 and 1 are used as the elements z and u respectively, but that 10 and 11 do not hold.

It will be an interesting and valuable game for the reader to work

out examples for which other groups, if not all, of these properties hold (see *27*). Before leaving him to this pastime, it will be convenient to introduce the word *field* which is used in a technical sense to designate any system in which all the above properties hold.

Definition IX. A *field* is a system consisting of a set of elements together with two operations on pairs of these elements of such a character that the properties 1–11 of *25* are all satisfied.

The enumeration of properties of the system of rational numbers in *25* can now be condensed in the statement that *the set of rational numbers with the operations addition and multiplication constitute a field.*

27. Playing on the walls.

1. Verify that $-\frac{2}{3} \div \frac{5}{6}$, $\frac{4}{5} + \frac{11}{7} - \frac{2}{15}$, $\frac{5}{3} \times \frac{4}{7}$, $\dfrac{\frac{2}{3} - \frac{1}{5}}{\frac{5}{3} - \frac{2}{5}}$ are rational numbers.

2. Prove that the set of negative integers is closed with respect to addition, but not with respect to multiplication.

3. Prove that the set of positive rational numbers is closed with respect to addition, multiplication, and division, but not with respect to subtraction.

4. Prove that the set consisting of the positive and negative integers and 0 is closed with respect to addition, subtraction, and multiplication, but not with respect to division.

5. Construct an addition table and a multiplication table for the set of numbers 0, 1, 2, 3, 4, 5, 6, 7, with the agreement that all numbers which exceed 7 are to be reduced by multiples of 8, until they leave a remainder that is in the set.

6. Test the set of numbers used in 5 for the field properties.

7. Show that the set of numbers $a + b\sqrt{3}$ in which a and b are positive integers is closed with respect to addition and multiplication. Which of the field properties are valid in this set?

8. Show that the set of numbers in 7 is not closed either with respect to subtraction or with respect to division. Enlarge the set so as to obtain closure (a) with respect to each of the inverse operations separately, (b) with respect to both.

9. Show that the set of numbers $3, 3^2, 3^3, 3^4, \ldots$ (i.e. the set of powers of 3 with positive integral exponents) is closed with respect to multiplication, but not with respect to the other rational operations.

10. Investigate the set of numbers $a + b\sqrt{5}$ in which a and b are rational numbers with respect to the rational operations and the field properties.

11. Investigate the set of numbers $a + bi$, in which a and b are rational numbers and $i^2 = -1$ with respect to the rational operations and the field properties.

12. Show that if the factor set is the set of numbers $a + b\sqrt{3}$, in which a and b are positive or negative integers or 0, then 1 and 2 are not prime numbers.

In questions 13–17, it will be supposed that a, b, c and d are elements of a *field:*

13. Prove that $a + (b + c) = a + (c + b) = b + (a + c)$
$$= b + (c + a) = c + (a + b) = c + (b + a).$$
Note. It follows from this proposition that the sum of 3 terms is independent of the order in which these terms appear.

14. Prove that $a(bc) = a(cb) = b(ac) = b(ca) = c(ab) = c(ba)$.

15. Prove that $(a + b) + (c + d) = a + (b + c) + d$ so that the symbol $a + b + c + d$ is unambiguous in meaning; prove also that its numerical significance is independent of the order of the terms.

16. Prove that $(ab)(cd) = a(bc)d$.

17. Prove that $(a + b)(c + d) = ac + ad + bc + bd$.

Remark. The reader will be able to extend indefinitely the number of propositions of the character of the last five and thus to keep himself occupied for a long time. They are special cases of the following general propositions: (1) the sum $a + b + c + \cdots + p$ of any number of elements of a field is independent of the order of these terms and of the groupings in which they are added together, (2) the product $abc \ldots p$ of any number of elements of a field is independent of the order of the factors and of the groupings in which they are multiplied together, (3) the product of any number of sums, each consisting of an arbitrary number of elements of a field, $(a_1+b_1+ \cdots +p_1)(a_2+b_2+ \cdots +p_2) \cdots (a_k+b_k+ \cdots +p_k)$ is equal to the sum of all the products obtainable by multiplying together any term of the first sum, any term of the second sum, etc.; thus the product would be equal to $a_1a_2 \cdots a_k + a_1a_2 \cdots b_k + \cdots + p_1p_2 \cdots p_k$. An excellent discussion of these matters may be found in J. H. Poincaré, *La Science et l'Hypothèse*, pp. 1–29; or in the book by Landau, *op. cit.*

CHAPTER IV

BREAKING THROUGH THE WALLS

It is the merest truism, evident at once to unsophisticated observation, that mathematics is a human invention. — P. W. Bridgman, *The Logic of Modern Physics*, p. 60.

28. Powers of numbers. One of the first new concepts to which the beginner in algebra is exposed, is that of the *powers of a number*. It grows out of the process of multiplication, just as the concept of the *multiples of a number* grows out of the process of addition. Everyone knows that $a + a + a + a + a$, the sum of 5 *terms* each of them equal to a, is called the product of a by 5 and is denoted by $5a$; and nearly everyone knows also that $a \times a \times a \times a \times a$, the product of 5 *factors* each of them equal to a, is called the *5th power* of a and is denoted by a^5; 5 is called the *exponent*, a the *base* of the power. This operation of raising a number to a power can be carried out without restriction within the set of natural numbers: 7^3, 3^7, 12^{17} are symbols which have definite meaning and to each of which corresponds a definite natural number. It can be carried out equally well if the base belongs to the set of rational numbers, so long as the exponent is restricted to the set of natural numbers: $(-3)^5$, $(\frac{2}{5})^4$, $(-\frac{3}{2})^2$ are also symbols with perfectly definite meaning, to each of which a definite rational number corresponds. What about powers in which both exponent and base can be chosen freely in the set of rational numbers? Many readers may know an answer to this question. For the present we shall not need it, because we are going to limit ourselves to powers whose base is an arbitrary rational number, but whose exponent is a natural number, i.e. to symbols of the form a^n, where a is an element of R, and n a natural number.[1]

29. Laws among powers. Of this set of numbers a few properties must be stressed:

1. Every power of a positive rational number is a positive

[1] We shall agree that if nothing is said to the contrary the letters n, m, l, k are to designate natural numbers.

rational number, i.e. the set of positive rational numbers is closed with respect to the process of "raising to a power" ("involution" is commonly used as an abbreviation for "raising to a power whose exponent is a natural number").

2. Every *odd* power of a negative rational number is a negative rational number, every *even* power is a positive rational number; i.e. the set of negative rational numbers is not closed with respect to the process of involution, but only with respect to the process of raising to an odd power.

3. The set of rational numbers is closed with respect to the process of involution.

The proofs of these statements can safely be left to the reader, but should not be neglected by him. We come now to some closure properties within the set of powers itself.

4. The set of powers of a single rational number a is closed with respect to multiplication, i.e. given two natural numbers n and m, there exists a single third natural number k such that $a^n \cdot a^m = a^k$.

The proof of this proposition follows at once from part (2) of the general proposition stated in the remark on page 46; and from it follows the further conclusion that $k = n + m$. Thus we obtain the well-known "law of exponents"

$$a^n \cdot a^m = a^{n+m}. \tag{4.1}$$

5. The set of powers of a single rational number a is closed with respect to involution, i.e. given two natural numbers n and m, there exists a single third natural number k such that $(a^n)^m = a^k$.

The proof of this proposition is a consequence of the preceding one; for it follows from (4.1) that $a^n \cdot a^n = a^{2n}$, hence that $a^n \cdot a^n \cdot a^n = a^{2n} \cdot a^n = a^{3n}$. Continuing in this way, we not only prove our proposition, but we discover at the same time that $k = nm$. Thus we obtain another law of exponents, viz.

$$(a^n)^m = a^{nm}. \tag{4.2}$$

6. If two powers of a single rational number a are given, a^n and a^m, such that $n > m$, there exists a single third natural number k such that $a^n \div a^m = a^k$, i.e. such that $a^m \cdot a^k = a^n$.

It follows from (4.1), that $k = n - m$; we have therefore obtained a third law of exponents, viz.

$$a^n \div a^m = a^{n-m}, \text{ provided } n > m. \tag{4.3}$$

We observe that the set of powers of a single rational number a is *not* closed under division. In accordance with the discussion in *24*, this gives us a suggestion for extension, for new adventure.

7. The nth power of the product (or quotient) of two rational numbers a and b is equal to the product (or quotient) of the nth powers of a and b, i.e.

$$(4.4) \qquad (ab)^n = a^n b^n \text{ and } \left(\frac{a}{b}\right)^n = \frac{a^n}{b^n}.$$

The first of these laws of exponents is also a consequence of part (2) of the remark made at the end of *27;* it suggests an immediate generalization to the form $(ab \ldots p)^n = a^n b^n \ldots p^n$. The second follows from the properties of rational numbers recalled in *23*.

The reader, if he has had the experience of a course in "freshman algebra" in high school, will recognize in the formulas (4.1), (4.2), (4.3) and (4.4), which express the content of propositions 4–7, old friends (or perhaps hated enemies, or bugbears) of his youth. A very large part of the formal work in elementary algebra courses consists in drill on these formulas. This drill is superfluous for any one who keeps the meaning of the symbols in mind. Mechanical application of formulas has no interest for us. We recall them here because they bring out some of the closure properties of the set of powers of rational numbers with natural numbers as exponents. Since these formulas follow from the field properties (see *25*), it should be clear that they are valid for the elements of any field.

30. A double-headed creature. The process of involution as we have considered it thus far introduces a new operation on rational numbers. Like addition and multiplication, it involves two numbers, viz. the base and the exponent. But it differs from these operations in two respects. In the first place, base and exponent do not have the same realms of eligibility (as did the two terms of a sum, and the two factors of a product). In the second place, it is not a commutative operation; base and exponent are not interchangeable with each other, as are the terms of a sum and the factors of a product. That is, while $a + b = b + a$, and $ab = ba$, we have not established the identity of a^n and n^a, not even if we restrict both a and n to the set of natural numbers. Indeed one example suffices to show that the commutative law does not hold between base and exponent of a power, even if both are selected

from the set of natural numbers; the fact that $2^3 = 8$, while $3^2 = 9$ is such an example. There appears here a lack of symmetry, such as did not occur in addition or multiplication. This asymmetry is of importance when we set out to obey the dictum of Jacobi (see p. 40) and to consider the inverse of involution. Instead of a single question, we get now two questions, viz.,

1. If b is a rational number and n a natural number, does there exist a rational number a such that $a^n = b$? and

2. If a and b are rational numbers, does there exist a natural number n, such that $a^n = b$?

Thus we find a situation, quite different from the one we ran into when we inverted addition and multiplication. Instead of a normal one-headed creature we have to do with a two-headed monster, which has two different ways of standing on his head. Or better, we have a relation between three elements, the base, the exponent and the power; and we must look at that relation from the point of view of each of them. When three individuals, or three groups, or three nations are to conclude a successful agreement, it has to be looked at by each of them, not only from his own point of view but also from the point of view of each of the others. This is what we shall now proceed to do with the relation $a^n = b$.

It turns out that the two inverse questions are quite different in character. Each of them will lead to an extension of our concept of numbers; both are far reaching, but they are of distinct kinds. The present chapter and the next will deal with the first question only. In Chapter VI, the two streams will then be brought together; not until then will we have acquired enough strength to master the double-headed creature.

31. A base camp. The road on which we are now starting has a considerable upward grade. We shall therefore begin by building a base of supplies in which we can take refuge. Instead of attacking question 1 in its full generality, we take first a more modest question, viz.,

3. If m is a positive integer does there always exist a positive rational number x such that $x^2 = m$?

The answer to this question will supply at the same time the answer to an apparently more general question, viz. if r is a positive rational number does there exist a rational number y such that $y^2 = r$? For, if $r = \dfrac{p}{q}$, we can always find a multiplier for numera-

tor and denominator which will change the denominator to a perfect square, i.e. we can reduce r to the form $\frac{m}{q_1^2}$, where m and q_1 are positive integers. The problem is then reduced to that of finding a rational number y such that $y^2q_1^2 = m$. But if we put $x = yq_1$, we come back to the equation $x^2 = m$. Since x is rational if y is, and conversely, we can limit ourselves without loss of generality, to a consideration of question 3. For example, the equation $y^2 = \frac{7}{12}$ is reduced successively to the forms $y^2 = \frac{21}{36}$, $36y^2 = 21$, $(6y)^2 = 21$, $x^2 = 21$, where $x = 6y$; the last question is of type 3.

It is easy to see that question 3 must be answered in the negative. A negative answer to a *general* question is substantiated, as soon as a single example has been adduced. (We have had an illustration of this fact on page 49.) In our present case, an example is furnished when we take $m = 2$. For if there were a positive rational number x such that $x^2 = 2$, this number could be written as the quotient of two relatively prime positive integers p and q (see p. 39) and we would then have $\left(\frac{p}{q}\right)^2 = 2$, or $p^2 = 2q^2$. That two such integers can not exist we see as follows: if p were odd, p^2 would be odd, it would not have a factor 2 and could therefore not equal $2q^2$; if p were even, it would have at least one factor 2, p^2 would have at least two factors 2, so that q^2 would have to have at least one factor 2, which could only happen if q were also even, in which case p and q would not be relatively prime. Hence our first attempt, viz. to find for an arbitrary positive integer m a positive rational number x such that $x^2 = m$, has met a decisive defeat at the very start at the hands of a single example. It is true that when $m = 1$, 4, 9, 16, 25, . . ., there exists every time a corresponding positive rational number x; but when $m = 2$ no such number exists and it would not be difficult to show that the same answer would have to be given for many other values of m such as 3, 5, 6, or 7. But a single contrary example is enough to upset a general law. Let us stop for a few observations on this statement.

In the first place, we can not fail to notice that the denial of validity to a general law on account of a single contrary example does not occur in many fields of human experience. Indeed the saying "the exception proves the rule" indicates a very different point of view, of which many illustrations are found in such fields as grammar, economics, sociology and some of the sciences. One

of the things which set mathematics apart from other human interests is that a mathematical proposition is not valid, unless it holds *without exceptions.* In the second place, we must remark that whereas one contrary example suffices to throw out a general law, one favorable example is not enough to establish it, and not two, nor three, nor a hundred or twenty thousand. And in this respect mathematics differs from the inductive sciences in which general laws are frequently established on the basis of observed facts. It is true that a careful worker in an inductive science will not be satisfied with only a few observations which confirm his "law." But in mathematics conclusions cannot be reached in that way at all; mathematics is a deductive science.

However, what are we to do with our equation $x^2 = m$? How are we going to turn our defeat into victory?

32. How much of the cake can we get? Let us return to the simple problem $x^2 = 2$. We know now that no rational number a exists which satisfies this equation. Not being able to get the whole cake, it is not inhuman for Johnny to see how large a part of it he can get. Taking this hint, we inquire how close we can come to 2 with the square of a rational number. Clearly, since $1^2 = 1$, and $2^2 = 4$, $a_0 = 1$ is as good an approximation as we can get in whole numbers; let us also record the closest approximation from above, $b_0 = 2$. There is a denumerable set of rational numbers between 1 and 2; the squares of some of them exceed 2, those of others are less than 2. In order to proceed systematically, we consider first the rational numbers 1.1, 1.2, 1.3, . . . 1.8 and 1.9. It is an easy matter to verify that

$$1.1^2 = 1.21,\ 1.2^2 = 1.44,\ 1.3^2 = 1.69,\ 1.4^2 = 1.96,$$
$$1.5^2 = 2.25,\ 1.6^2 = 2.56,\ 1.7^2 = 2.89,\ 1.8^2 = 3.24,\ 1.9^2 = 3.61,$$

so that the first four have squares less than 2, while the other five have squares greater than 2. The best approximation from below is $a_1 = 1.4$, the best approximation from above is $b_1 = 1.5$. In a_1 we have a rational number whose square differs from 2 by .04. Is this as close as we can get? It is if we limit ourselves to rational numbers of the form $1 + \frac{p}{10}$, where p is an integer less than 10. But between 1.4 and 1.5 there is again a denumerable infinity of rational numbers; from them we select the set 1.41, 1.42, . . . 1.48, 1.49. We find at the cost of a little work that $1.41^2 = 1.9881$,

and $1.42^2 = 2.0164$. Hence we have obtained the approximations $a_2 = 1.41$ and $b_2 = 1.42$, whose squares differ from 2 by .0119 and .0164 respectively. And now every one will understand that this process may be carried on another step, and another step, and another step, indefinitely. Thus we would find successively

$a_3 = 1.414$, $a_3^2 = 1.999396$; $b_3 = 1.415$, $b_3^2 = 2.002225$.
$a_4 = 1.4142$, $a_4^2 = 1.99996164$; $b_4 = 1.4143$, $b_4^2 = 2.00024449$.
$a_5 = 1.41421$, $a_5^2 = 1.9999899241$; $b_5 = 1.41422$, $b_5^2 = 2.0000182084$.

If we continue this process far enough we can obtain rational numbers a_k and b_k whose squares differ from 2 by as small an amount as may be desired. It will be worth our while to prove this last statement, even though the reader may hardly feel the need for a proof; for this proof will bring out some useful points.

We observe (1) that $b_0 - a_0 = 1$, $b_1 - a_1 = \frac{1}{10}$, $b_2 - a_2 = \frac{1}{10^2}$, $b_3 - a_3 = \frac{1}{10^3}$, etc., and, in general, we will have $b_k - a_k = \frac{1}{10^k}$.
(2) That $3 > b_0 > b_1 > b_2 > \ldots$, while $a_0 < a_1 < a_2 < \ldots < 2$; hence, in general, $b_k + a_k < 5$.
(3) That $a_0^2 < 2 < b_0^2$, $a_1^2 < 2 < b_1^2$, $a_2^2 < 2 < b_2^2$, $a_3^2 < 2 < b_3^2$ etc. and, in general, $a_k^2 < 2 < b_k^2$.

Consequently,

$$2 - a_k^2 < b_k^2 - a_k^2 = (b_k + a_k)(b_k - a_k) < \frac{5}{10^k};$$

and also,

$$b_k^2 - 2 < b_k^2 - a_k^2 = (b_k + a_k)(b_k - a_k) < \frac{5}{10^k}.$$

If k is a large enough integer, $\frac{5}{10^k}$ will be less than any desired amount. For instance, if the desired amount is $\frac{1}{7{,}834{,}529}$, we have but to take $k = 8$, since

$$\frac{5}{10^8} = \frac{5}{100{,}000{,}000} = \frac{1}{20{,}000{,}000} < \frac{1}{7{,}834{,}529};$$ and so on.

Thus, as a partial compensation for the defeat we suffered in *31*, we have now obtained the following result:

Theorem VI. There exist rational numbers whose squares differ from 2 by an amount as small as may be desired.

Remark 1. The reader should observe that the phrase "if we continue this process" is presupposed in the statements made under (1), (2) and (3) above.

Remark 2. The process as far as we have indicated it does not show us a very convenient way of actually determining the rational numbers a_0, a_1, a_2, a_3 etc. and b_0, b_1, b_2, b_3 etc. The numbers a_0, a_1, a_2, a_3 etc. are usually found by the scheme called "Square root extraction" in most books on arithmetic and elementary algebra. The reader may remember this scheme. We are not concerned with it here, but rather with the fact that such rational numbers exist; the invention of a numerical scheme for determining them is a separate problem, of a technical character. Numerical schemes of this sort are frequently called *algorithms.*

The reader who has followed the discussion which led to Theorem V should have no difficulty in understanding that two sets of rational numbers a_0, a_1, a_2, a_3 etc., and b_0, b_1, b_2, b_3 etc. can be determined such that (1) $b_k - a_k = \frac{1}{10^k}$, (2) $a_0 \leqq a_1 \leqq a_2 \leqq a_3 \leqq \ldots$, and $b_0 \geqq b_1 \geqq b_2 \geqq b_3 \geqq \ldots$, (3) $a_k^2 \leqq m \leqq b_k^2$, where m is any positive integer. Thus we can give the following answer to the question put at the beginning of *31.*

Theorem VIa. If m is a positive integer there does not always exist a rational number x such that $x^2 = m$; but there always exist rational numbers whose squares differ from m by an amount as small as may be desired.

The form of statement here used applies in first instance when m is not a perfect square. But it is applicable as well if m is a perfect square. For, if $m = k^2$, and k is an integer, then for the rational numbers whose existence is asserted we can take k, no matter how small the amount that "is desired." If we have the equation $x^2 = 25$, we could use successively the following rational numbers:

$$a_0 = 5,\ b_0 = 6;\ a_1 = 5.0,\ b_1 = 5.1;\ a_2 = 5.00,\ b_2 = 5.01;$$
$$a_3 = 5.000,\ b_3 = 5.001.$$

In this way we would still have sets of rational numbers, which satisfy the conditions (1), (2) and (3) mentioned just before Theorem VI*a*. We see moreover why we used the combination symbols $\leqq$ and $\geqq$, instead of the simple symbols $<$ and $>$. Could other sets of rational numbers be used here?

33. Work for the young people. But we can get a good deal more out of this process. Let us consider the question of determining a rational number x such that $x^3 = 9$. If there were such a rational number, it could be put in the form $\frac{p}{q}$, where p and q are relatively prime positive integers, so that we would have $p^3 = 9q^3$. Now it is clear that no such integers p and q can exist. For if p has no factor 3, its cube does not have the factor 3 and hence it can not equal $9q^3$; and if p does have a factor 3, its cube must have three factors 3, so that q^3 would have to have at least one such factor which could not happen unless q had at least one factor 3, in which case it would not be relatively prime to p. So far, we have again a purely negative result. But, since $2^3 = 8$ and $3^3 = 27$, $a_0 = 2$ and $b_0 = 3$ are approximations in the sense that their cubes differ from 9 by 1 and 18 respectively, b_0 being a rather poor approximation. If we investigate the rational numbers 2.1, 2.2, . . . 2.9 between 2 and 3, we find that the cubes of all of them exceed 9; for $(2.1)^3 = 9.261$. Hence we take $a_1 = 2.0$ and $b_1 = 2.1$. The next step consists in considering the rational numbers 2.01, 2.02, . . . 2.09.

At the cost of some labor, which can be much reduced by the use of a computing machine (or of our young brothers and sisters, for whom it is valuable practice), we find that $(2.08)^3 = 8.998912$, and $(2.09)^3 = 9.129329$. We take therefore $a_2 = 2.08$ and $b_2 = 2.09$. Continuing the process we obtain

$$\begin{aligned}
a_3 &= 2.080, & a_3^3 &= 8.998912000;\\
a_4 &= 2.0800, & a_4^3 &= 8.998912000000;\\
a_5 &= 2.08008, & a_5^3 &= 8.999850375936512;\\[1ex]
b_3 &= 2.081, & b_3^3 &= 9.011897441;\\
b_4 &= 2.0801, & b_4^3 &= 9.000209982401;\\
b_5 &= 2.08009, & b_5^3 &= 9.000080178544729.
\end{aligned}$$

The cube of a_5 differs from 9 by slightly more than .0001, that of b_5 by slightly less than that amount. We observe that the cubes of these successive rational numbers a_0, a_1, a_2, a_3, a_4, a_5 and b_0, b_1, b_2, b_3, b_4, b_5 do not tend towards 9 as rapidly as the squares of the numbers determined in 32 approached 2. Nevertheless we can prove that if the process here outlined is carried on, we obtain rational numbers a_k and b_k whose cubes differ from 9 by as small

an amount as may be desired. The reader may here feel the need of proof a little more than he did at the corresponding point in *32;* the same method can be used and we will benefit from having used it once before (as a practice game) although its use may have seemed superfluous at the earlier point. We have in a very similar way

(1) $b_k - a_k = \frac{1}{10^k}$;

(2) $b_k \leqq 3$, $a_k < 3$, and hence $a_k^2 + a_k b_k + b_k^2 < 27$; and

(3) $a_k^3 < 9 < b_k^3$, and hence $9 - a_k^3 < b_k^3 - a_k^3$.[1]

But now a slight change occurs. For, instead of $b_k^2 - a_k^2$, we have to consider $b_k^3 - a_k^3$. Now the reader may remember (and if he does not remember, he can verify this by a simple multiplication; see also p. 35 and p.205), that

$$b_k^3 - a_k^3 = (b_k - a_k)(b_k^2 + a_k b_k + a_k^2).$$

We can therefore proceed as follows:

$$9 - a_k^3 < b_k^3 - a_k^3 = (b_k - a_k)(b_k^2 + a_k b_k + a_k^2) < \frac{27}{10^k}.$$

By taking k large enough, $\frac{27}{10^k}$ can be made smaller than any amount that may be desired. And in the same way, we show that $b_k^3 - 9$ can be made smaller than any desired amount. We record our result as follows:

Theorem VII. There exists no rational number whose cube equals 9, but there exist rational numbers whose cubes differ from 9 by any amount as small as may be desired.[2]

Having gone this far in our study of the inverse of involution, we can now supply at least a much bigger part of the answer to question 1 in *30* than we have given before, as follows:

Theorem VIII. If m is a positive integer and n a natural number, there does not always exist a rational number x, such that $x^n = m$; but there do exist positive rational numbers whose nth powers differ from m by as small an amount as may be desired.

Proof. As in the special cases which we have treated in detail, if there exists no integer x for which $x^n = m$, we can find two consecutive positive integers a_0 and b_0 such that $a_0^n < m < b_0^n$. Then

[1] Observe that in this case $a_1 = a_0$, and $a_2 = a_3 = a_4$; also that the laws on inequalities of Chapter II (see p. 28) come in very conveniently.

[2] An algorithm for the determination of these rational numbers is found in books on algebra, e.g. Fine, *A College Algebra.*

we can find among the rational numbers a_0, $a_0 + \frac{1}{10}$, $a_0 + \frac{2}{10}, \cdots$ $a_0 + \frac{9}{10}$, b_0, two, let us call them a_1 and b_1, such that $b_1 - a_1 = \frac{1}{10}$ and $a_1^n < m < b_1^n$. Next, from among the rational numbers a_1, $a_1 + \frac{1}{100}$, $a_1 + \frac{2}{100}$, $\cdots$ $a_1 + \frac{9}{100}$, b_1, we find two successive ones, a_2 and b_2, such that $b_2 - a_2 = \frac{1}{100}$ and $a_2^n < m < b_2^n$. Continuing in this way we build up two sets of rational numbers $a_1, a_2, a_3, \ldots$ and $b_1, b_2, b_3, \ldots$ such that (1) $b_k - a_k = \frac{1}{10^k}$, (2) $a_k^n < m < b_k^n$, (3) $a_0 \leqq a_1 \leqq a_2 \leqq a_3 \leqq \ldots$, and $\ldots b_3 \leqq b_2 \leqq b_1 \leqq b_0$.

To complete the proof we need to make use of another bit of elementary algebra (the forgetful reader will know by this time where to refresh his memory), viz. that for every natural number n, $b_k^n - a_k^n = (b_k - a_k)(b_k^{n-1} + b_k^{n-2}a_k + b_k^{n-3}a_k^2 + \cdots + b_k a_k^{n-2} + a_k^{n-1})$; we notice that the second factor contains exactly n terms. With this preparation, we obtain the desired result easily, viz.,

$$m - a_k^n < b_k^n - a_k^n = (b_k - a_k)(b_k^{n-1} + b_k^{n-1}a_k + \cdots + a_k^{n-1})$$
$$\leqq \frac{1}{10^k} \cdot nb_0^{n-1}; \text{ and similarly } b_k^n - m < \frac{nb_0^{n-1}}{10^k}.$$

Since nb_0^{n-1} is a fixed number, we can, no matter how large this fixed number may be, take k so large that $\frac{nb_0^{n-1}}{10^k}$ is less than any amount that may be desired; thus our proof is completed.

Let us consider as an example the equation $a^7 = 472$. We find easily that $a_0 = 2$ and $b_0 = 3$; hence $nb_0^{n-1} = 7 \cdot 3^6 = 5{,}103$. Suppose now some one "desired" .000000001; how large must we take k? Since $nb_0^{n-1} < 10^4$, it follows that $\frac{nb_0^{n-1}}{10^k} < \frac{10^4}{10^k} = \frac{1}{10^{k-4}}$ (laws of exponents, see p. 48). In order to make this less than the desired amount which is $\frac{1}{10^9}$, we have to see to it that $k - 4 > 9$, i.e. $k > 13$. Actually to carry out the calculations, at least by the process we have used before, would not be an enviable task; it would not even be advisable to impose it upon the younger members of our families. But, we repeat, we are not worrying at all about the algorithm. The essential thing is that we have established the existence of rational numbers whose nth powers approximate an arbitrary positive integer as closely as may be

desired. By using the argument made on page 51, we obtain from Theorem VIII the following further result:

Corollary. If r is a positive rational number and n a natural number, there does not always exist a rational number y, such that $y^n = r$; but there do exist positive rational numbers whose nth powers differ from r by as small an amount as may be desired.

Let us consider, for example, the equation $y^3 = \frac{7}{4}$. We put it in the form $y^3 = \frac{14}{8}$ or $(2y)^3 = 14$. Putting $x = 2y$, this leads to $x^3 = 14$, an equation of the type we have considered above. Suppose now that $a_0, a_1, a_2, a_3, \ldots a_k, \ldots$ and $b_0, b_1, b_2, b_3, \ldots b_k, \ldots$ are the sets of positive rational numbers which are determined for this problem, so that

$$14 - a_k^3 \leqq \frac{3b_0^2}{10^k}, \quad \text{and} \quad b_k^3 - 14 \leqq \frac{3b_0^2}{10^k}.$$

Then $\frac{7}{4} - \left(\frac{a_k}{2}\right)^3 \leqq \frac{3}{8} \cdot \frac{b_0^2}{10^k} < \frac{3b_0^2}{10^k}$ and $\left(\frac{b_k}{2}\right)^3 - \frac{7}{4} \leqq \frac{3}{8} \cdot \frac{b_0^2}{10^k} < \frac{3b_0^2}{10^k}$; hence the sets of positive rational numbers $\frac{a_0}{2}, \frac{a_1}{2}, \frac{a_2}{2}, \frac{a_3}{2}, \ldots \frac{a_k}{2}, \ldots$ and $\frac{b_0}{2}, \frac{b_1}{2}, \frac{b_2}{2}, \frac{b_3}{2}, \ldots \frac{b_k}{2}, \ldots$ serve the same purpose for the equation $y^3 = \frac{7}{4}$.

34. Breaking through the walls. We have not yet obtained a complete answer to question 1 of *30*. It will have to be postponed until a later chapter; and the answer to question 2 will not come until after that. But let us try to find out what extension of the system of rational numbers is suggested by the answer we have obtained so far.

We begin with the definition of a word we have already used, and which will facilitate our further discussion.

Definition X. A *sequence* is an ordered arrangement of the elements of a denumerable set (see p. 18).[1]

We recall that a denumerable set is a set which is equivalent to the set of natural numbers. If we actually establish a 1–1 correspondence of which this equivalence asserts the existence, the set is ordered. We have then a 1st element, a 2nd element, a 3rd element, and so on; a sequence appears then in the form $e_1, e_2, e_3, \ldots$. It is important to observe at this point that a denumerable

[1] An ordered arrangement of a finite set is frequently referred to as a *finite sequence*.

set may be ordered in several ways, i.e. that it may give rise to different sequences. For example, the set of natural numbers may be arranged, apart from the natural order, in the following sequences:

$$2, 1, 4, 3, 6, 5, 8, 7, \ldots 2n, (2n-1), \ldots$$

$$5, 3, 1, 2, 4, 10, 8, 6, 9, 7, 15, 13, 11, 14, 12, \ldots;$$

and in many others. The set of proper fractions may be ordered in the sequence

$$\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \tfrac{2}{3}, \tfrac{1}{5}, \tfrac{1}{6}, \tfrac{2}{5}, \tfrac{3}{4}, \tfrac{1}{7}, \tfrac{3}{5}, \tfrac{1}{8}, \tfrac{2}{7}, \tfrac{4}{5}, \ldots$$

suggested in *14*. But it can also be ordered as follows:

$$\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{2}{3}, \tfrac{1}{4}, \tfrac{3}{4}, \tfrac{1}{5}, \tfrac{2}{5}, \tfrac{3}{5}, \tfrac{4}{5}, \tfrac{1}{6}, \tfrac{5}{6}, \ldots.$$

So much for the concept *sequence*. The content of the discussion leading up to Theorem VIII and to its corollary can now be stated as follows: If r is a positive rational number and n a natural number, there exist a non-decreasing sequence of positive rational numbers $a_0, a_1, a_2, \ldots a_k, \ldots$ and a non-increasing sequence of positive rational numbers $b_0, b_1, b_2, \ldots b_k, \ldots$, such that every element of the first sequence is less than every element of the second, and such that a_k^n and b_k^n differ from r by at most $\dfrac{nb_0^{n-1}}{10^k}$.

Of course there are many rational numbers besides those which occur in these two sequences. Now we throw in with the sequence $a_0, a_1, a_2, \ldots$ all those rational numbers which are *less* than any element of this sequence; and we throw in with the sequence $b_0, b_1, b_2, \ldots$ all those rational numbers which are *greater* than any element of it. Let us designate the two sets of rational numbers obtained in this way by A and B. Then it follows

1. That every rational number belongs either to A or to B. For, if a rational number s does not belong to A it is greater than any element of the sequence $a_0, a_1, a_2, a_3, \ldots$; since the difference $b_k - a_k$ can be made less than any amount, by taking k large enough, s must surely exceed some element b_k and therefore belong to B. A similar argument shows that if s does not belong to B, it must belong to A.

2. Every element of A is less than every element of B. For if s belongs to A it is surely less than some element a_k, which itself is less than every element of the set b_0, b_1, etc. and therefore surely less than every element of B; hence s is less than every element of B.

The sets A and B constitute therefore a special kind of division of the entire class R of rational numbers (see p. 42) into two mutually exclusive parts. Such a division is technically called a *cut*.

Definition XI. A *cut* of the system R of rational numbers is a division of R into two sets A and B, such that every element of R belongs either to A or to B, and such that every element of A is less than every element of B.

We shall denote such a cut by the symbol (A, B).

This concept was introduced into mathematics by the German mathematician R. Dedekind (1831–1916), a contemporary and friend of Georg Cantor; it has played a very important part in clarifying the fundamental basis of the science. It supplies us with an extension of the set R which serves our present purpose. By means of it we break through the wall in which the restriction to the rational operations has kept us hemmed in. We enter now into the extended domain, which consists of all *cuts* in the set of rational numbers.

The discussion of *32–34* shows us that, if r is a positive rational number, and n a natural number, there exists a cut (A, B) of R such that for any positive element a of A, $a^n \leqq r$, for any element b of B, $b^n > r$, and such that no matter how small the rational number e is, there are elements a and b such that $b - a < e$. This is true also if there exists a rational number r_1 for which $r_1{}^n = r$. For with the rational number r_1 we can associate the cut (A, B) in which A consists of all the rational numbers which are *less than or equal to* r_1, and B of all the rational numbers which *exceed* r_1.

Thus the set of cuts in R contains an element corresponding to every element of R itself; besides it contains elements to which correspond no element of R. We observe now that if a cut corresponds to a rational number r, then r is the largest element of the set A. But if (A, B) is a cut to which no element of R corresponds, then A has no largest element and B has no least element. Conversely, to every cut (A, B) in which A possesses a largest element r, we make correspond the rational number r. Thus we have established a 1–1 correspondence between the set of rational numbers R and those cuts (A, B) in which A has a largest element; we shall call such cuts *rational cuts*. Let us denote by C the set of all cuts in R (omitting those in which B has a least element) and by C_R the set of rational cuts. Then *the sets R and C_R are equivalent sets* (see p. 15).

Examples.

1. To the equation $x^3 = 8$ corresponds the cut in which A consists of all rational numbers less than or equal to 2, and B of all rational numbers greater than 2.

2. To the equation $x^2 = 2$ corresponds the cut in which A consists of all rational numbers which are less than or equal to any of the rational numbers $a_0 = 1$, $a_1 = 1.4$, $a_2 = 1.41$, etc. which are determined by any algorithm for the extraction of a square root, while B consists of all rational numbers greater than or equal to any of the rational numbers $b_0 = 2$, $b_1 = 1.5$, $b_2 = 1.42$, etc.; in general, $b_k = a_k + \frac{1}{10^k}$ (compare pp. 52, 53).

35. To improve the acquaintanceship.

1. Prove that the set of powers of the rational number a whose exponents are natural numbers is closed under multiplication, but not under addition, subtraction or division.

2. Develop an extended set of powers of the rational number a in order to obtain closure with respect to division as well as with respect to multiplication. *Hint.* The extension has to remove the restriction $n > m$ from (4.3) on page 48.

3. Explain the meaning given to powers of a rational number a with integral exponents, positive, negative, or zero in the extended set of the preceding problem.

4. Use formulas (4.1) to (4.3) to prove:
(a) $a^n \cdot a^m \cdot a^p = a^{n+m+p}$, a being a rational number and n, m, p natural numbers; (b) $[(a^n)^n]^n = a^{n^3}$; (c) $(a^n)^3 = a^{3n}$; (d) $(a^n)^m \neq a^{n^m}$.

5. Reduce to a power of a:

$$\text{(a)} \left(\frac{a^5 \cdot a^4}{a^2}\right)^3; \text{ (b) } \frac{(a^5)^3 \cdot (a^4)^5}{(a^3)^4}; \text{ (c) } \left[\frac{(a^5)^4}{(a^3)^5}\right]^4 \div \left[\frac{(a^4)^3}{(a^5)^4}\right]^3.$$

6. Prove the generalization of (4.4), mentioned on page 49.

7. Use the extension of (4.4) to show, without using (4.1) or (4.2) that $(a^n)^m = (a^m)^n$.

8. Reduce to the product of a power of a by a power of b:

$$\text{(a) } \frac{(a^3)^4 \cdot (b^5)^2}{(a^2b^3)^3}; \text{ (b) } \frac{(a^2b^3)^4 \cdot (a^3b^4)^5}{(a^4b^5)^2 \cdot (a^5b^2)^3}; \text{ (c) } \left[\left(\frac{a^5}{b^2}\right)^4 \div \left(\frac{a^3}{b^4}\right)^5\right]^3.$$

9. Determine a rational number whose square differs from 8 by less than .0001; settle in advance how many decimal places have to occur in this rational number.

10. Prove that no rational number exists whose cube is equal to $\frac{7}{5}$. Determine a rational number whose cube differs from $\frac{7}{5}$ by less than .01; how many decimal places must it contain?

11. Exhibit the cuts in R which correspond to the rational numbers $-\frac{5}{7}, \frac{12}{5}$.

12. Show that the equation $x^2 = 2$ gives rise to a second cut in R besides the one determined by the sequences of which the initial elements are exhibited on page 53.

36. Operations on cuts, and isomorphism. The suggestion with which the preceding section closes indicates that while for every rational number there is but a single cut in R, there may be more than one cut corresponding to an equation of the form $x^n = r$, where r is a positive element of R and n a natural number. In fact, it should be clear that, if n is *even* and the equation leads to the non-decreasing sequence of positive rational numbers $a_0, a_1, a_2, a_3, \ldots$ and the non-increasing sequence of positive rational numbers $b_0, b_1, b_2, b_3, \ldots$, then the non-decreasing sequence of negative rational numbers $-b_0, -b_1, -b_2, -b_3, \ldots$ and the non-increasing sequence of negative rational numbers $-a_0, -a_1, -a_2, -a_3, \ldots$ also have the property that nth powers of their elements approximate r to within any desired degree of accuracy. For, since n is even $(-a_k)^n = a_k{}^n$, and $(-b_k)^n = b_k{}^n$; so that the nth powers of the elements of the sequences of negative numbers differ from r by the same amounts as the nth powers of the corresponding elements of the sequences of positive members. The reader should have no difficulty in understanding that the sequence $-a_0, -a_1, -a_2, \ldots$ is non-increasing, if the sequence $a_0, a_1, a_2, \ldots$ is non-decreasing; similarly that to the non-increasing sequence $b_0, b_1, b_2, \ldots$ corresponds the non-decreasing sequence $-b_0, -b_1, -b_2, \ldots$. Hence to such an equation corresponds besides the cut (A, B), a second cut which we shall represent by the symbol $(-B, -A)$ in which $-B$ represents the set of all rational numbers less than or equal to any element of the non-decreasing sequence $-b_0, -b_1, -b_2, \ldots$, and $-A$ the set of all rational numbers greater than or equal to any element of the non-increasing sequence $-a_0, -a_1, -a_2, \ldots$. It is natural to call the cut $(-B, -A)$ the negative of the cut (A, B), and to write $(-B, -A) = -(A, B)$.[1]

[1] If the cut (A, B) is a rational cut which corresponds to a positive rational number r, a slight change has to be made in the definition of the cut $(-B, -A)$; in that case $-A$ is the set of all rational numbers greater than $-r$, and $-B$ the set of all rational numbers less than *or equal* to r.

We have thus carried over to cuts, in a very simple manner one of the operations, viz. that of taking the negative, i.e. of multiplying by -1, which are applied to the rational numbers themselves. A more extended study of the theory which is here sketched would lead us to definitions for the comparison, the addition, subtraction, multiplication and division of cuts. For instance, we would call the cut (A_2, B_2) greater than the cut (A_1, B_1) if there were an element common to the sets A_2 and B_1; we would then show that the fundamental rule 1 for inequalities (see p. 28) holds. We would distinguish between *positive cuts* (A, B), of which the set A contains positive rational numbers, *negative cuts* (A, B), of which the set B contains negative rational numbers and a *zero cut* (A, B) for which the set A consists of zero and all the negative rational numbers and B of all the positive rational numbers. We would add two cuts (A_1, B_1) and (A_2, B_2) by forming a cut (A, B) in which A consists of all the rational numbers obtainable by adding any element of A_1 to any element of A_2, and B of all the sums of an element of B_1 and one of B_2. It would have to be shown that in this way we would indeed obtain a cut, i.e. a division of R which satisfies Definition XI. This more extended study we can not undertake here; we have too many other things to do. The reader will find it discussed quite fully in the first part of Dedekind's *Essays on Numbers* (translated by W. W. Beman),[1] and in the books by Landau and Perron referred to on page 39. We shall indicate some of the results of this study.

1. C, the set of cuts in R, has all the properties of a field with respect to addition and multiplication as defined for cuts, (see *25*).

2. If c is a positive or zero cut and n a natural number, there exists a positive cut x such that $x^n = c$.

3. If the rational cuts (A_1, B_1) and (A_2, B_2) correspond to the rational numbers r_1 and r_2 respectively, then the cut $(A_1, B_1) + (A_2, B_2)$ corresponds to the rational number $r_1 + r_2$, the cuts $(A_1, B_1) - (A_2, B_2)$, $(A_1, B_1) \cdot (A_2, B_2)$ and $(A_1, B_1) \div (A_2, B_2)$ to the rational numbers $r_1 - r_2$, r_1r_2 and $\frac{r_1}{r_2}$ respectively; with the proviso that for the division r_2 shall be different from zero.

4. The rules for inequalities (see p. 28) are also valid in the set C. Moreover, if r_1 and r_2 are rational numbers such that $r_1 < r_2$ and

[1] This little volume furnishes very interesting and useful reading.

(A_1, B_1) and (A_2, B_2) the corresponding rational cuts, then $(A_1, B_1) < (A_2, B_2)$; if $r_1 = r_2$, then $(A_1, B_1) = (A_2, B_2)$, and if $r_1 > r_2$, then $(A_1, B_1) > (A_2, B_2)$.

Of these properties of the set C the first one includes closure with respect to the rational operations; the second one asserts that the operation involution can be inverted at least in one way, within the set C_+ of positive cuts, i.e. the set C_+ is closed with respect to the first inverse of involution. The third and fourth properties show that the equivalence of the sets R and C_R (see p. 60) is maintained under the rational operations, and that relations of order among elements of R are valid without change for the corresponding elements of C_R. The sets R and C_R are thus seen to be more intimately related than could be inferred from the mere fact of their equivalence. In a certain sense, they are indistinguishable from each other without being identical; indistinguishable, at least, in so far as the results of the rational operations and ordering by the relations greater than, equal to, less than are concerned. It will be convenient to introduce a technical term for this more intimate correspondence before continuing the discussion.

Definition XII. Two equivalent sets, together with one or more pairs of corresponding relations between elements in each of them are called *isomorphic* (i.e. of the same shape or structure) if the equivalence between the sets is maintained with respect to the relations.

Let us look at this a little more explicitly. Take the first set to be A_1, its elements $a_1, b_1, c_1, \ldots$ (the set does not have to be denumerable); suppose that there are relations $\rho_1(a_1, b_1)$, $\sigma_1(a_1, b_1, c_1)$, $\lambda_1(a_1, b_1, c_1, d_1)$ etc. which hold between these elements. For instance if A were a set of numbers or of cuts, $\rho_1(a_1, b_1)$ might mean $a_1 > b_1$; $\sigma_1(a_1, b_1, c_1)$ might mean $a_1b_1 = c_1$, $\lambda_1(a_1, b_1, c_1, d_1)$ might mean $a_1 + b_1 = c_1d_1$. Take now a second set A_2 with elements a_2, $b_2, c_2, \ldots$ and relations $\rho_2, \sigma_2, \lambda_2$; it is to be understood that if ρ_1 is a relation between *two* elements A_1, then ρ_2 is a relation between *two* elements of A_2; that σ_1 and σ_2, λ_1 and λ_2 are in this same sense pairs of corresponding relations. We consider now the *system* $\{A_1, \rho_1, \sigma_1, \lambda_1, \ldots\}$ consisting of the set A_1 and the relations ρ_1, σ_1, λ_1, etc. . . .; and similarly we have the *system* $\{A_2, \rho_2, \sigma_2, \lambda_2, \ldots\}$. When we say that the systems $\{A_1, \rho_1, \sigma_1, \lambda_1, \ldots\}$ and $\{A_2, \rho_2, \sigma_2, \lambda_2, \ldots\}$ are isomorphic we mean that (1) there is a 1–1 correspondence between the elements of A_1 and A_2, (2) there is a

1–1 correspondence between the relations ρ_1, σ_1, λ_1, etc. and ρ_2, σ_2, λ_2, etc., (3) if we suppose that a_1, b_1, c_1, etc. correspond to a_2, b_2, c_2, etc. respectively, and ρ_1, σ_1, λ_1, etc. to ρ_2, σ_2, λ_2, etc. then, whenever any of the relations $\rho_1(a_1, b_1)$, $\sigma_1(a_1, b_1, c_1)$, $\lambda_1(a_1, b_1, c_1, d_1)$ hold, the relations $\rho_2(a_2, b_2)$, $\sigma_2(a_2, b_2, c_2)$, $\lambda_2(a_2, b_2, c_2, d_2)$ hold also, and vice versa.

Definition XIII. A 1–1 correspondence between two systems $\{A_1, \rho_1, \sigma_1, \lambda_1, \ldots\}$ and $\{A_2, \rho_2, \sigma_2, \lambda_2, \ldots\}$, each consisting of a set and relations between its elements, which shows that the two systems are isomorphic is called an *isomorphism.*

Thus the results from the theory of cuts which were mentioned under 3 and 4 on page 63 can be stated by saying that the system consisting of the set of rational numbers R with the rational operations and order relations among them is isomorphic with the system consisting of the set C_R of rational cuts and the rational operations and order relations as defined for the elements of that set.

37. Moving into larger quarters. The answer to our problem should now be clear to the reader. The discussion of the preceding pages points directly to the abandonment of the set of rational numbers R and to substitution for it of the set C consisting of all cuts in R. For, in the first place, C contains a subset C_R isomorphic with R; and in the second place the inversion of involution is possible at least within the set C_+ of positive cuts.[1] But, before we can bring the discussion to its conclusion, we must gain a somewhat deeper insight into the significance of isomorphic systems.

When two *systems* are isomorphic, either is a complete image of the other. Although the elements as well as the relations between them may be totally different in the one system from what they are in the other, everything that occurs in the one system has its counterpart in the other; the two systems are related as two languages spoken by two peoples with identical background and experiences. If either people were suddenly and completely to forget their own language during sleep and to acquire instead an equivalent knowledge of the other, they would have no way of discovering the change. During recent years, in discussions of the theory of relativity, it has often been suggested that if all dimensions of our bodies and of the objects in our world of experience, including the measuring instruments, were suddenly to be reduced

[1] In the next chapter we shall see how the exclusion of negative cuts can be done away with.

in size to one half of their present values, there would be no way in which this change could be discovered. Old and new measurements together with old and new relations between measurements would be isomorphic. In just such a way the change from the system consisting of the set R and the rational operations within that system to the system consisting of the set C_R and the rational operations within it is not discernible by experiences within either system alone.

The isomorphism between R and C_R is of great value when we abandon R and move over into the larger set C. For, all rational operations performed upon elements of C_R have their counterpart in operations in R familiar from childhood. Let us then use the familiar symbols for rational numbers to designate the corresponding elements of C_R and operate with them as we have always done (it is here assumed that we have always done it correctly); our conclusions will then be valid within C_R. In other words, "plus que ça change, plus que c'est la même chose." In final analysis nothing has been changed, except in our conceptions; but, furthermore, the set C_R is imbedded in the larger set C, within which we can move about more freely than we could in R (see *31*, *32*, *33*).

It was in order to remove this limitation upon our operations that the search for an extension was made. What has been accomplished has been an enlargement of our field without the sacrifice of any earlier results. Thus real progress has been made, not a mere change from one system to another, but an extension in which is contained all that was useful in the old. In technical language, we have obtained a new domain a *proper* part of which is isomorphic with the old domain. This might not be a bad formula for measuring progress in other human undertakings.

38. Irrational numbers are also real. We repeat that this extension from R to C has so far given us only a partial answer to the first question discussed in *30*. We will however see later on (see Chapter VI) that it enables us to deal not only with the first inversion of the process of involution (see p. 50, question 1), but also with the second inversion (see p. 50, question 2), provided we limit ourselves to positive rational numbers. In fact, the set C contains a great deal more than is needed for the restricted problem discussed in this chapter.

The reader will have had sufficient introduction by now to the way in which mathematics creates its own worlds and to its en-

largement of the domain of primitive "common sense," not to be surprised when he is told that the elements c of C are also called numbers, in fact *real numbers.*

Definition XIV. A *real* number is a cut in the set of rational numbers.

We have here the definition of real numbers which was promised on page 4 and in several other places after that.

The positive cut which provides the answer to the equation $x^2 = 2$ is denoted by the symbol $\sqrt{2}$.[1] Similarly the positive cut which solves the equation $x^3 = 7$ is denoted by the symbol $\sqrt[3]{7}$; and, in general, if r is a positive rational number and n a natural number, the symbol $\sqrt[n]{r}$ designates the positive cut c, for which $c^n = r$. Thus $\sqrt{2}$, $\sqrt[3]{7}$, $\sqrt[n]{r}$ are all *positive real numbers.* The rational cuts, i.e. the elements of the set C_R, are called *rational real numbers* in order to distinguish them from the rational numbers; we must always remember that the sets R and C_R are indeed isomorphic, but not identical. Cuts which are not rational cuts are called *irrational numbers.*

Real numbers suffice for describing lengths, areas, volumes, etc. which often can not be obtained by direct measurement (compare p. 38). For direct simple measurements the rational numbers are sufficient; but not for indirect measurements. If we want to determine the length of the hypothenuse of a right triangle, whose right sides have been found by direct measurement to be each 1 unit in length, this can not be done by direct measurement. But the real number $\sqrt{2}$ measures this hypothenuse; similarly $\sqrt[3]{7}$ serves to measure in length units the edge of a cube whose volume is equal to 7 volume units; and so forth.

Among the other numbers contained in the set of real numbers is the "number" π, which the reader will remember as the measurement in area units of the area of a circle whose radius is 1 length unit. It will be worth while to sketch a method by which a cut for this number may be constructed. We begin by inscribing in the circle of unit radius regular polygons of 3, 6, 12, 24, 48 sides and so forth. By methods of elementary geometry we can determine real numbers which measure the areas of these successive polygons; let

[1] A complete treatment must include a proof of the fact that the product of this cut, which is denoted by $\sqrt{2}$, by itself is actually equal to the rational cut which corresponds to 2.

us denote them by a_0, a_1, a_2, a_3, a_4, etc.[1] Then we circumscribe about the circle polygons of 3, 6, 12, 24, 48 sides and so forth, and determine the real numbers which measure their areas; let the results be b_0, b_1, b_2, b_3, b_4, etc.[2] Thus we have obtained a non-decreasing (in fact increasing) sequence of real numbers, $a_0 < a_1 < a_2 < a_3 < \ldots < a_k < \ldots$, and a non-increasing (actually decreasing) sequence of real numbers $b_0 > b_1 > b_2 > b_3 > \ldots > b_k > \ldots$. We form now a cut (A, B) in R by putting in A all the rational numbers which are less than or equal to any one of the real numbers a_k, and in B all the rational numbers which are greater than or equal to any one of the real numbers b_k. This cut is the real number π.[3]

[1] Books on elementary geometry show (a) that the area of the regular inscribed polygon of n sides is given by

$$\frac{1}{4} n s_n \sqrt{4 - s_n^2}$$

where s_n is the length of one side of the polygon, (b) that $s_3 = \sqrt{3}$, and (c) that

$$s_{2n} = \sqrt{2 - \sqrt{4 - s_n^2}}.$$

Using these facts, we find

$$a_0 = \frac{3\sqrt{3}}{4},\ a_1 = \frac{3\sqrt{3}}{2},\ a_2 = 3,\ a_3 = 6\sqrt{2 - \sqrt{3}},\ a_4 = 12\sqrt{2 - \sqrt{2 + \sqrt{3}}},$$

$$a_5 = 24\sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{3}}}},\ \text{etc.}$$

[2] We learn in elementary geometry (a) that the area of the regular circumscribed polygon of n sides is given by $\frac{nS_n}{2}$, where S_n is the length of one side, (b) that $S_3 = 2\sqrt{3}$, and (c) that

$$S_{2n} = \frac{2\sqrt{4 + S_n^2} - 4}{S_n}.$$

From these formulas we obtain

$$b_0 = 3\sqrt{3},\ b_1 = 2\sqrt{3},\ b_2 = 12(2 - \sqrt{3}),\ b_3 = 24(2\sqrt{2 + \sqrt{3}} - \sqrt{3} - 2),\ \text{etc.}$$

[3] To show completely that this division of the set of rational numbers is a cut, we have to prove that by taking k large enough $b_k - a_k$ can be made less than any amount no matter how small. This can be done as follows: (a) Every a_k is certainly less than $b_0 = 3\sqrt{3}$, (b) it is shown in elementary geometry books that

$$S_k = \frac{2s_k}{\sqrt{4 - s_k^2}};$$

(c) hence

$$b_k - a_k = \frac{kS_k}{2} - ks_k\frac{\sqrt{4 - s_k^2}}{4} = \frac{k}{2}\left[2s_k\frac{\sqrt{4 - s_k^2}}{4 - s_k^2} - s_k\frac{\sqrt{4 - s_k^2}}{2}\right]$$

$$= ks_k\frac{\sqrt{4 - s_k^2}}{2}\left[\frac{2}{4 - s_k^2} - \frac{1}{2}\right] = \frac{a_k s_k^2}{4 - s_k^2} < \frac{b_0 s_k^2}{4 - s_k^2} = \frac{3\sqrt{3}s_k^2}{4 - s_k^2};$$

(d) but $s_k \leqq s_3 = \sqrt{3}$, so that $4 - s_k^2 > 4 - 3 = 1$ and therefore $b_k - a_k < 3\sqrt{3}s_k^2$. Since by taking k large enough s_k can be made arbitrarily small, it follows that $b_k - a_k$

In the course of our further work, we hope to become better acquainted with the content of the set of real numbers. For the present we rest content with the following result which summarizes the work of this chapter.

Theorem IX. The set of real numbers has all the properties of a field. Moreover, if b is any positive real number and n a natural number, there always exists a unique positive real number a such that $a^n = b$.

39. More recreation.

1. Prove that the set of natural numbers is isomorphic with the set of positive integers with respect to addition and multiplication (compare p. 16).

2. Prove that the set of natural numbers is isomorphic with the set of negative integers with respect to addition.

3. Prove that the set of integers, positive, negative or zero, is isomorphic with the set of rational numbers whose denominator is 1, with respect to the rational operations, excepting division (compare p. 39, footnote).

4. Prove that the set of even natural numbers is isomorphic with the set of natural numbers with respect to addition. (Is this also true for the set of odd natural numbers?)

5. Prove that the system consisting of the natural numbers and addition is isomorphic with the system consisting of the positive integral powers of the positive rational number a and multiplication.

6. Connecting with *35*, 3 show how an isomorphism may be established between the system consisting of integers, positive, negative, or zero and addition, and the integral powers of the rational number a and multiplication.

7. Establish a cut for the real number π by using inscribed and circumscribed regular polygons of 4, 8, 16, 32 sides and so forth.

8. There are two weak spots in the argument made in footnote 3 on page 68; one is caused by an inadequate notation, the other by a tacit assumption. Locate them; clarify them as much as possible.[1]

9. Prove, by use of Theorem IV (see p. 29) that the set of real numbers is not denumerable.

can be made less than any desired amount. If we wish $b_k - a_k$ to be less than .0001, we have to see to it that $s_k^2 < \frac{.0001}{3\sqrt{3}}$; this will surely be the case if $s_k < \frac{.01}{3}$.

[1] There is no escape from vagueness here. The weak spots were left in the argument not in order to provide an exercise but because avoiding them seemed only possible at the cost of introducing more details than was deemed desirable. Once they are there, attention should be called to them.

10. Deduce from the preceding problem the conclusion that the set of irrational numbers is not denumerable; obtain in this way an illustration of the answer in *10*, 5 (see p. 17).

11. Prove that if a system S_1 is isomorphic with a system S_2, and S_2 is isomorphic with a system S_3, then S_1 is isomorphic with S_3.

12. Prove that if a system S_1 is isomorphic with a system S_2, and a system T_1 with a system T_2, then the system obtained by adjoining T_1 to S_1 is isomorphic with the system obtained by adjoining T_2 to S_2.

CHAPTER V

TO GREATER FREEDOM

L'esprit n'use de sa faculté créatrice que quand l'expérience lui en impose la nécessité. — Poincaré, *La Science et l'Hypothèse*, p. 43. (The mind only uses its creative powers when experience presents it with the necessity for such use.)

40. Vectors and scalars. In the sense explained in *38*, the real numbers satisfy the needs of measurement of lengths, weights, areas and volumes and so forth. But there are many other magnitudes which occur in human experience. When we wish to nail up a box, we apply force, and we know that it is not only the amount of force that matters but also the direction. When a pulling force is exerted on a railway carriage, on a cart, on a boat, or on a float, the direction of the pull is of great importance. It would be silly to quote more examples of something with which everybody is familiar; having been started off in this direction, the reader will readily carry on, unless he is acted upon by a force to change his direction. Force is a magnitude which has *amount* and *direction*, it is a *directed magnitude*. There are many other directed magnitudes; among them *velocity* and *acceleration* readily come to mind. And a common property of all these directed magnitudes is that their direction as well as their amount is unrestricted. The name vector has been given to such magnitudes. Magnitudes which are completely given by an amount only are called *scalars*.

Definition XV. A *vector* is a magnitude which is arbitrary in its direction as well as in its amount. A *scalar* is a magnitude whose amount is arbitrary but whose direction is specific.

Real numbers are not adequate to the measurement of vectors. It is true that we can distinguish by means of real numbers between two opposite directions, like North and South, hotter and colder, increase and decrease; for this purpose the $+$ and $-$ signs are adequate. But as soon as we depart from such a mere variation of *sense* and come to a freely varying direction, they are no longer sufficient. How are we going to remedy this defect and adjust our set of numbers to this new need?

41. The arithmetic of points. In the first place we introduce a geometric representation of vectors. At its foundation lies an assumption. This is not surprising. Whenever we dig up a structure which has been "logically" put together, we shall find assumptions at the foundation. The assumption made here is of special interest, because it underlies the possibility of an interchange between analysis and geometry.

Assumption. There is an isomorphism between the system of real numbers together with the rational operations upon them and the order relations between them, and the system consisting of all the points on a directed straight line, together with the rational operations on these points and the order relations between them.

This statement requires some further explanations.

1. A directed straight line is a straight line upon which a point

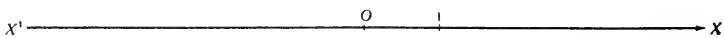


FIG. 7

has been marked off, called the origin O, and on which a direction has been indicated (see Fig. 7). It is usual, although not essential, to have the indicated direction on a horizontal line go to the right from the origin.

2. By marking off an arbitrary point as the unit point, we can establish a 1–1 correspondence between the integers, positive, negative and zero, and certain points on the line. Points corresponding to positive integers are obtained by laying off the distance $O1$ to the *right* of O as many times as there are units in the integers; the point O corresponds to 0, and points which correspond to negative integers are obtained by laying off the distance $O1$ to the *left* of O as many times as there are units in the integer.

3. To determine a point corresponding to the rational number $\frac{p}{q}$, we proceed as follows: through O we draw a second line, $Y'OY$, (see Fig. 8), on which we indicate a direction, and a unit which may be but need not be equal to the unit on OX. On OY we mark off points P and Q, corresponding to the *integers* p and q, by the method explained in 2. If Q is joined to 1, and through P a line is drawn parallel to $Q1$, this line will meet OX in a point A; this is the point which we make correspond to the rational number $\frac{p}{q}$.

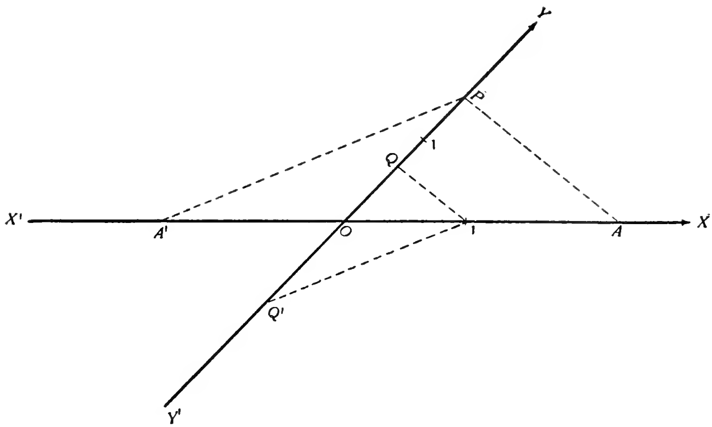


FIG. 8

The reader will easily see that if $\frac{p}{q}$ is positive, so that p and q may both be taken as positive, the point A will come to the right of O; whereas if $\frac{p}{q}$ is negative, so that p may be taken to be positive and q negative, the corresponding point, will (as A' in Fig. 8) lie to the left of O.

4. So far there is nothing very mysterious about our assumption. But we have only assigned points so far to the rational real numbers. When it comes to the irrational numbers we do not have such clear sailing if we want to indicate a construction for the corresponding point. It is true that we could give one for numbers of the form $\sqrt{r}$, where r is a positive rational number, but when it comes to $\sqrt[3]{7}$, $\sqrt[5]{17}$ or π, —! Failing a construction, the assumption becomes indeed an *assumption*. Our geometric intuition can be brought into play to make it palatable, by making us see that if we have points corresponding to all the rational numbers of the sequences a_0, a_1, a_2, etc. and b_0, b_1, b_2, etc. (those in the one sequence non-decreasing, those of the other non-increasing and such that the difference $b_k - a_k$ falls below every amount, as k increases, see Ch. IV), then it is not unreasonable to suppose that to this cut corresponds some point on the line; but nevertheless it remains an assumption to say that the set of real numbers and the set of points on a directed line are equivalent.

This is the content of the first part of the assumption. It is really not new for us; for we have tacitly assumed it and used it when we introduced the "coördinates" of a point (see pp. 21, 25). Indeed it is common practice to talk about real numbers as *being* points on a line, which they certainly *are* not (unless the poor words *are*, *is* etc. receive quite extraordinary interpretations). We also talk about the points to the right of the origin as *positive* points, and about those to the left as *negative* points, when to be exact we should say "points, corresponding to positive real numbers" etc. Such exactness is superfluous and pedantic; but we should have in mind the real meaning of the shorter phrase which we use and bring it into the foreground when the occasion demands it.

5. Order relations among points A, B, C, etc. on $X'OX$ are expressed in the form: A is to the left of B, to the right of C (see Fig. 9).

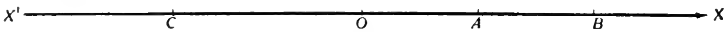


FIG. 9

The assumption maintains that if the points A, B, C, etc. correspond respectively to the real numbers a, b, c, etc., then whenever $a < b$, A is to the left of B; whenever $a > c$, A is to the right of C; and when $a = d$, A coincides with D. In accordance with it, the symbols $<$, $>$ and $=$ are frequently used in relation to points with the meanings "to the left of," "to the right of," and "coincident with."

6. Associated with the rational operations on real numbers we have corresponding constructions on points, as follows:

If a segment equal to OB in amount and direction is laid off beginning at A, and if its endpoint is C, we say that C is obtained by "adding" B to A.

The assumption says that if a, b and c are the real numbers which correspond to the points A, B and C respectively, and if C is obtained by "adding" B to A, then $c = a + b$, and conversely.

To "subtract" B from A, we lay off a segment equal in amount and direction to BO, beginning at A; if it ends at C, then C corresponds to $a - b$.

The "multiplication" and "division" of points is carried out as follows:

An arbitrary line $Y'OY$ is drawn, as in 3, with a direction and a unit (see Fig. 10). Through B we draw a line parallel to $11'$, meeting $Y'OY$ in B'; through B' we draw a line parallel to $A1'$,

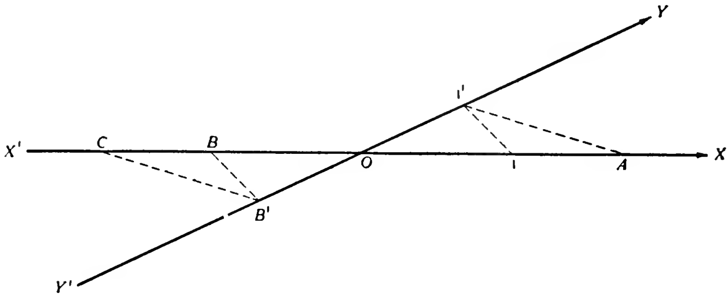


FIG. 10

meeting $X'OX$ in C. The point C is the "product" of the points A and B.

With the same machinery as before we draw AA' and BB' parallel to $11'$, and $A'C$ parallel to $B'1$. The point C so obtained is the quotient of A by B, see Fig. 11.

It is an interesting exercise for the reader to verify that, if A and B are points of like sign, their product and quotient are positive

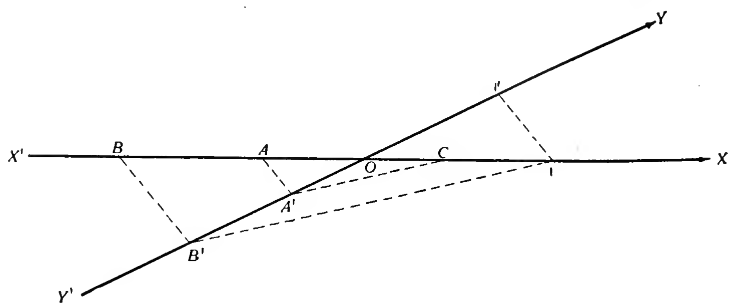


FIG. 11

points, whereas, if these points have opposite signs, the product and quotient are negative points. Another valuable exercise consists in verifying that the field properties are satisfied by the points of a line and the operations on them which have been specified above. An exhaustive discussion of these matters can be found in the important book by the contemporary German mathematician

D. Hilbert, *The Foundations of Geometry*, pp. 46–50. The reader should now have a clear understanding of the assumption stated on page 72.

42. Number pairs represent vectors. The point O together with an arbitrary point A on the line $X'OX$ furnishes an adequate geometric representation of a vector whose amount is given by the length of OA measured in terms of the unit specified on the line, and whose direction coincides either with the positive or the negative direction specified on that line.

Let us now return to a consideration of general vectors, or directed magnitudes. Their amount is measurable by means of real numbers. How about their direction? Direction is essentially a geometrical concept; to specify it we need geometrical machinery. The possible directions that can be realized depend upon the space of our geometry. We have already seen that mathematicians do not shrink from the study of spaces of a more general character than "common sense" would admit. Developments in modern physics have led to the consideration of vectors in such general spaces. For our present purpose we shall limit ourselves however to vectors in a plane; they are called *plane vectors*.

The special case mentioned in the first paragraph suggests a simple way to represent plane vectors geometrically. Taking O

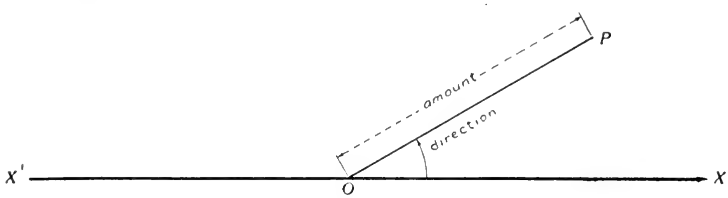


FIG. 12

as the initial point in every case, the representation of such a vector is completely determined as soon as the terminal point P has been indicated; this point P may come anywhere in the plane. In this representation, the amount of the vector is specified by the measure of the *undirected* segment OP in terms of the unit on $X'OX$, and its direction by the angle XOP (see Fig. 12). By this method only the specification of a point P in the plane is required for the representation of a plane vector. For every plane vector we have a point P in the plane; conversely, every point P in the plane

represents a plane vector. (Which vector is represented by the point O?)

It should now be a simple matter to determine a set of *numbers* adequate to represent plane vectors. For we know that the set of points P in a plane is equivalent to the set of pairs of real numbers (a, b), the coördinates of the points P (compare 13). The vector being represented by the directed line OP, and this representation being determined by the position of the terminal point P, we obtain as a natural *numerical* representation of the plane vector the pair

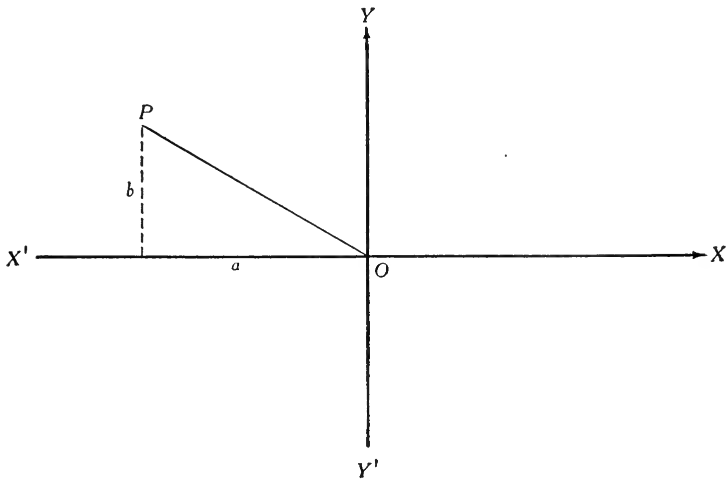


FIG. 13

of real numbers which are the coördinates of P with respect to the coördinate axes obtained by adjoining the line $Y'OY$ to the line $X'OX$, perpendicular to it (see Fig. 13). To every plane vector corresponds a pair of real numbers; and to every such pair there corresponds a plane vector. In particular to the number pairs of the form $(a, 0)$, representing points on the line $X'OX$, correspond the vectors whose direction is that of OX, or of OX'. The need for a numerical representation of plane vectors has thus led us to an extension from the set C of real numbers to the set of *pairs of real numbers;* we shall designate the latter set by C_2.

43. A program. But we are not yet out of the woods; we have a few more matters to settle, as follows:

1. If C_2 is to be adequate to represent plane vectors, it must be possible to add, subtract, multiply and divide the elements of C_2 in such a way as to have the results correspond to the sum, difference, product and quotient of vectors, as suggested, at least in part, by their physical significance; i.e. it does not suffice for C_2 to be equivalent to the set of vectors — the two systems {plane vectors, rational operations} and {C_2, rational operations} must be *isomorphic*.

2. If the transition from C to C_2 is to be a true extension, the latter set must contain a proper part not merely equivalent to C but isomorphic with it with respect to the rational operations.

The problem before us then is to determine definitions for the sum, difference, product and quotient of two pairs (a_1, b_1) and (a_2, b_2) which are adequate to these needs.

44. A new way of adding and subtracting. It follows from the manner in which the number pairs were introduced that (a_1, b_1) and (a_2, b_2) are to be considered as equal only if they correspond to the same point, i.e. if $a_1 = a_2$ and $b_1 = b_2$.

Definition XVI. The pairs (a_1, b_1) and (a_2, b_2) are called *equal* if and only if $a_1 = a_2$ and $b_1 = b_2$.

For addition and subtraction we can be guided by the addition of vectors. I hope that every reader has heard of the *parallelogram-*

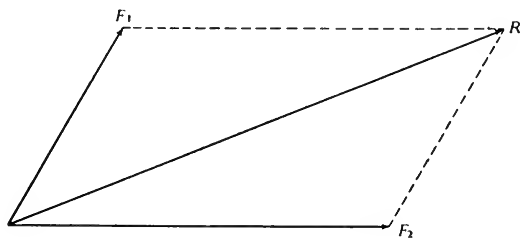


FIG. 14

law of Newton, which states in a general and abstract way what everybody has experienced, viz.: If two forces F_1 and F_2 act on the same point, (see Fig. 14), the resultant is represented by the diagonal of the parallelogram, of which two adjacent sides represent these forces. As concrete examples of this "law," we have the effect on a boat of the pulls exerted on it by means of two ropes, the effect on a ball of two simultaneous kicks from different directions. As a special case we have the effect on a bone caused by two dogs

who pull at it in opposite directions, etc. The same process is used in defining the sum of two vectors. Hence, the sum of the number pairs (a_1, b_1) and (a_2, b_2) represented by the points P_1 and P_2 respectively (see Fig. 15) must be the number pair (a, b) which is represented by the point P, the fourth vertex of the parallelogram, determined by O, P_1 and P_2. To determine a and b, we observe that $OS = a$,[1] and $SP = b$; also that the triangles OP_1R_1 and

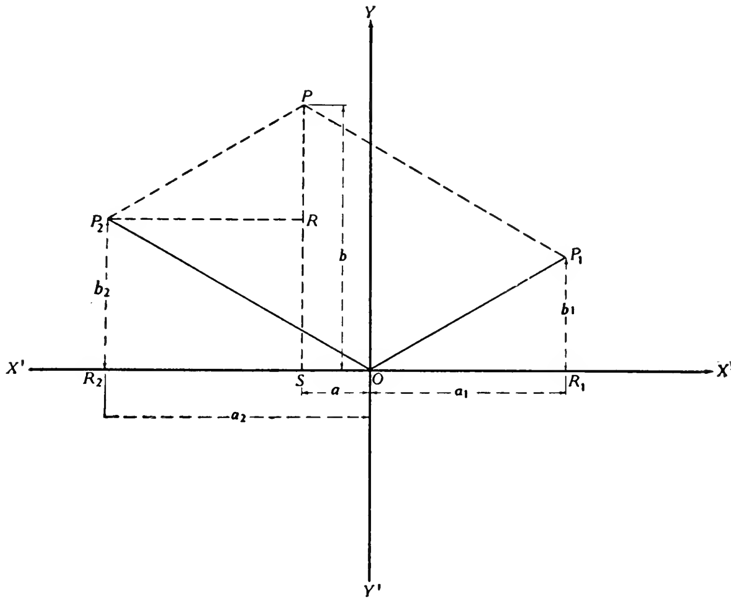


FIG. 15

P_2PR are congruent, because all their angles are equal and $P_2P = OP_1$. Consequently $P_2R = OR_1 = a_1$ and $RP = R_1P_1 = b_1$. Therefore $a = OS = OR_2 + R_2S = OR_2 + P_2R = a_2 + a_1$, and $b = SP = SR + RP = R_2P_2 + RP = b_2 + b_1$. We lay down therefore the following definition:

Definition XVII. The sum of the pairs (a_1, b_1) and (a_2, b_2) is the pair $(a_1 + a_2, b_1 + b_2)$:

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2). \tag{5.1}$$

[1] It is essential to write OS and not SO; in the present case a is a negative number; caution is necessary throughout to observe the directions of segments.

The reader will do well to carry through the construction of Fig. 15 for various positions of the points P_1 and P_2.

Since the set of real numbers possesses all the field properties (see p. 69) it follows that $a_2 + a_1 = a_1 + a_2$, and $b_2 + b_1 = b_1 + b_2$, so that the addition of pairs as defined above is commutative (see p. 42), i.e. $(a_1, b_1) + (a_2, b_2) = (a_2, b_2) + (a_1, b_1)$. This is as it should be, for the requirement 2 of *43* carries with it that the set C_2 must have all the field properties. We shall repeatedly make use of this fact in the sequel.

To subtract (a_1, b_1) from (a_2, b_2) means to find a pair (a, b) such that $(a_1, b_1) + (a, b) = (a_2, b_2)$ (man muss immer umkehren, see p. 40). This means, on the basis of Definitions XVII and XVI, that $a_1 + a = a_2$, and $b_1 + b = b_2$, and hence that

$$a = a_2 - a_1, \quad \text{and} \quad b = b_2 - b_1.$$

It is therefore not a new definition, but a consequence of the earlier definitions and of the definition of subtraction as the inverse of addition, that

$$(5.2) \qquad (a_2, b_2) - (a_1, b_1) = (a_2 - a_1, b_2 - b_1).$$

It is convenient to introduce at this point the negative of a number pair (the element n_a of *25*, 10) as follows:

Definition XVIII. The negative of a pair is a new pair, whose elements are the negatives of the corresponding elements of the given pair:

$$(5.3) \qquad -(a, b) = (-a, -b).$$

It follows from (5.1), (5.2) and (5.3) that

$$(a_2, b_2) - (a_1, b_1) = (a_2, b_2) + [-(a_1, b_1)];$$

hence by the use of negative number pairs, subtraction becomes again a special case of addition. Moreover the zero-element in C_2 (see *25*, 8) is the pair (0, 0).

45. Completing the constitution. To work out multiplication in C_2 in accordance with the requirements of *43*, we draw once more upon the concrete example of plane vectors furnished by forces.

1. The multiplication of a force by a positive number is effected by leaving its direction unchanged and multiplying its amount by that number. For instance, to multiply a pull by 2 means to keep on pulling in the same direction, but twice as hard.

Carrying this principle over to general vectors and hence to number pairs leads, when we remember that by 2 of 43, a scalar should correspond to a pair of the form $(a, 0)$, to the following result: If $a > 0$, then $(a_1, b_1) \cdot (a, 0) = (a_1a, b_1a)$. For the direc-

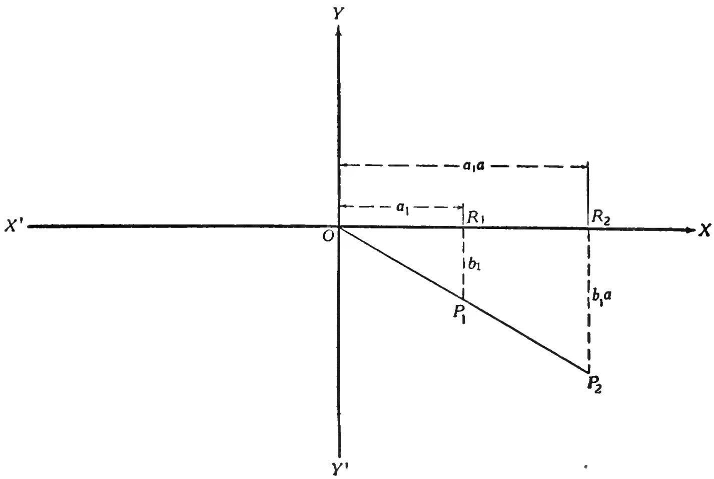


FIG. 16

tions of the vectors represented by P_1 and P_2 are the same, while it follows from the similarity of $\triangle OP_1R_1$ and $\triangle OP_2R_2$ that $OP_2 : OP_1 = b_1a : b_1$.

Hence $$OP_2 = OP_1 \cdot \frac{b_1a}{b_1} = OP_1 \cdot a.$$

2. The multiplication of a force by -1 is effected by leaving its amount unchanged and reversing its direction.

By applying this principle, we find that

$$(a_1, b_1) \cdot (-1, 0) = (-a_1, -b_1).$$

Combination of these two formulas shows that formula (5.4) is valid also if a is negative; and clearly it is valid as well if $a = 0$. Since moreover, as has already been observed, C_2 must have all the properties of a field, multiplication of pairs must be commutative; we can conclude that

$$(5.4) \qquad (a_1, b_1) \cdot (a, 0) = (a, 0) \cdot (a_1, b_1) = (a_1a, b_1a).$$

3. In particular, it follows from (5.4) that

(5.5) $$(a, 0) \cdot (0, 1) = (0, a).$$

Now $(a, 0)$ corresponds to a vector from O to a positive or negative point on the X-axis, and $(0, a)$ to a vector of the same amount to a positive or negative point on the Y-axis; this last vector can be obtained from the first by rotating it in counter-clockwise sense through an angle of $90°$ about O. This leads us *to interpret the effect of multiplication by $(0, 1)$ throughout* as a rotation through $90°$ about O in counter-clockwise sense. If carried out on a vector from O to a point on the positive (negative) Y-axis, there results a vector of the same amount from O to a point on the negative (positive) X-axis; i.e.

(5.6) $$(0, a) \cdot (0, 1) = (-a, 0).$$

4. A combination of the results obtained thus far leads us, under application of the field properties, quickly to an adequate definition of the product of two arbitrary pairs. From (5.1), (5.5), (5.6) and the distributive law follows:

$$(a_1, b_1)(0, 1) = [(a_1, 0) + (0, b_1)]\,(0, 1) = (0, a_1) + (-b_1, 0) = (-b_1, a_1).$$

Hence, by use of the associative law and of (5.4):

$$(a_1, b_1)(0, b_2) = (a_1, b_1)[(0, 1)(b_2, 0)] = [(a_1, b_1)(0, 1)](b_2, 0) = (-b_1, a_1)(b_2, 0) = (-b_2b_1, b_2a_1).$$

Finally,

$$(a_1, b_1)(a_2, b_2) = (a_1, b_1)\,[(a_2, 0) + (0, b_2)] = (a_1a_2, b_1a_2) + (-b_1b_2, a_1b_2) = (a_1a_2 - b_1b_2, a_1b_2 + a_2b_1).$$

This last result is readily seen to contain as special cases all the preceding formulas obtained in this section; they are therefore all satisfied if it is adopted as the definition of multiplication. Accordingly, we put down:

Definition XIX. The product of two pairs (a_1, b_1) and (a_2, b_2) is the pair whose first element is $a_1a_2 - b_1b_2$ and whose second element is $a_1b_2 + a_2b_1$:

(5.7) $$(a_1, b_1)(a_2, b_2) = (a_1a_2 - b_1b_2, a_1b_2 + a_2b_1).$$

It is readily verified that with this definition, multiplication in C_2 satisfies all the field properties which relate to multiplication.

Division remains to be considered. We base its interpretation

on its definition as the inverse of multiplication. Hence to divide (a_1, b_1) into (a_2, b_2) means the determination of a pair (a, b) such that

$$(a_1, b_1)(a, b) = (a_2, b_2).$$

It follows from Definitions XVI and XIX that a and b must then satisfy the following conditions: $a_1a - b_1b = a_2$ and $b_1a + a_1b = b_2$. From this we derive, by eliminating first b and then a, that

$$a = \frac{a_1a_2 + b_1b_2}{a_1^2 + b_1^2}, \quad \text{and} \quad b = \frac{a_1b_2 - a_2b_1}{a_1^2 + b_1^2}.$$

Hence we obtain, not as a new definition, but as a consequence of earlier definitions, that

$$(5.71) \qquad \frac{(a_2, b_2)}{(a_1, b_1)} = \left(\frac{a_1a_2 + b_1b_2}{a_1^2 + b_1^2}, \frac{a_1b_2 - a_2b_1}{a_1^2 + b_1^2}\right).$$

As a special case of (5.71), we find for the reciprocal (the r_a of 25, 11) of a pair (a_1, b_1) the formula:

$$(5.72) \qquad \frac{(1, 0)}{(a_1, b_1)} = \left(\frac{a_1}{a_1^2 + b_1^2}, \frac{-b_1}{a_1^2 + b_1^2}\right).$$

Combining (5.7), (5.71) and (5.72) shows that

$$(a_2, b_2) \cdot \frac{(1, 0)}{(a_1, b_1)} = \left(\frac{a_2a_1 + b_1b_2}{a_1^2 + b_1^2}, \frac{a_1b_2 - a_2b_1}{a_1^2 + b_1^2}\right) = \frac{(a_2, b_2)}{(a_1, b_1)},$$

so that division is a special case of multiplication; the two operations, as should be the case in a field, have melted into one.

For the reader is left the interesting and simple task of completing the verification of the field properties as a consequence of Definitions XVII, XVIII and XIX.

The system consisting of the set C_2 and the operations addition and multiplication as here defined is the new system with which we are now going to be occupied; it is called the system of Complex Numbers.

Definition XX. The system of complex numbers is the set C_2 of pairs of real numbers (a, b) in which addition and multiplication are defined by means of Definitions XVII and XIX respectively.

46. There is nothing imaginary in "Imaginary Numbers." The definitions of the preceding section have been framed principally in accordance with the second requirement of *43* for the simple

reason that significance can not easily be given to the product of two vectors. We shall return to this question presently (see *49*, p. 89). Let us see what the consequences are of the fulfillment of this second requirement.

The set of complex numbers of the form $(a, 0)$ together with the rational operations upon them form a system that is isomorphic with the system $\{C, \text{rational operations}\}$. As we have seen in Chapter IV this means that certainly as far as rational operations are concerned the set of pairs $(a, 0)$ is indistinguishable from, although not identical with, the set of real numbers. There can then arise no confusion if we use the symbol a instead of the symbol $(a, 0)$; we can operate with a as if it were a real number and re-interpret it as a complex number at any stage of the operations.

The set C_2 of complex numbers contains another proper part of special interest, viz. the set of pairs of the form $(0, b)$. For it follows from Definition XVII that

$$(5.8) \qquad (0, b_1) + (0, b_2) = (0, b_1 + b_2),$$

so that if this set of pairs is put into 1–1 correspondence with the set C by making the pair $(0, b)$ and the real number b correspond to each other, this correspondence is maintained under addition. How about multiplication? We find that as a consequence of Definition XIX,

$$(5.9) \qquad (0, b_1) \cdot (0, b_2) = (-\, b_1 b_2, 0).$$

Clearly then, the correspondence between the set of pairs $(0, b)$ and the set of real numbers b is *not* maintained under multiplication. But, if we set up a correspondence between the set of pairs $(0, b)$ and the set of symbols ib, in which b is a real number, and adjoin to it the correspondence between the set of pairs $(a, 0)$ and C, this correspondence is maintained under multiplication provided we agree to operate with i as with an ordinary algebraic symbol which has the special property $i^2 = -1$. For to the relation $(0, b_1) \cdot (0, b_2) = (-\, b_1 b_2, 0)$ would then correspond $ib_1 \cdot ib_2 = -\, b_1 b_2$, which is indeed fulfilled if the symbol i has the property $i^2 = -1$. Moreover to $(a_1, 0) \cdot (a_2, 0) = (a_1 a_2, 0)$ would correspond $a_1 \cdot a_2 = a_1 a_2$, and to $(a, 0) \cdot (0, b) = (0, ab)$ would correspond $a \cdot ib = iab$; both of these relations hold.

Therefore the system which consists of the sets of pairs of the form $(a, 0)$ and $(0, b)$ together with multiplication is isomorphic

with the sets of numbers a and ib, under the agreement that $i^2 = -1$.

Here we have arrived, probably from an unexpected road, at a familiar point; for the "numbers" ib are seen to be nothing else than the "normal numbers" of page 35. The reader will now perhaps find some justification for the name "normal number"; it was used because to these numbers correspond vectors which are "normal" (i.e. perpendicular) to the vectors which correspond to real numbers. He may also know that the numbers ib are usually called "imaginary numbers." Indeed the name "normal numbers" is an innovation which is proposed in the hope that it will divest these perfectly innocent numbers of the awe-inspiring mysteriousness which has always clung to them. The name "imaginary" was undoubtedly a natural outcome of the way in which these numbers made their appearance in mathematics. It will be useful to read in this connection what the repeatedly quoted books of Dantzig and Conant have to say on this point; or to consult one of the numerous books on the history of mathematics.[1] The developments of the present section have shown that these numbers can be thought of as answering the need for the representation in one set, of vectors whose directions are mutually perpendicular. Since the long word "perpendicular" is commonly replaced throughout mathematics by "normal," the term "normal number" is suggested. Some may object that it would be unfortunate thus to remove the element of romantic mysteriousness which has always clung to the "imaginary numbers." Let them not fear. There is a great deal of romance and of mystery left in mathematics for him who has the courage to go in search of new adventures. We do not need to maintain artificially the pretense that there is mystery where there was only ignorance. But enough on this point; we will continue to use the term "normal numbers," to refer to numbers of the form ib, in which b is a real number, and i an ordinary algebraic symbol with the special property $i^2 = -1$. If at any time this should cause us worry we have but to remember that this ib is a convenient symbol for the number pair $(0, b)$ in which b is a real number.

One aspect of the question remains to be settled. While we have seen that the systems $\{(a, 0), (0, b);$ multiplication$\}$ and $\{a, ib;$

[1] See, e.g. F. Cajori, *A History of Mathematics*, pp. 146, 166, or V. Sanford, *A Short History of Mathematics*, pp. 185–87.

multiplication} are isomorphic, this isomorphism does not include addition. Indeed neither the set consisting of pairs $(a, 0)$ and $(0, b)$, nor that consisting of the numbers a and ib is closed under addition. This defect is remedied easily. Since $(a, 0) + (0, b) = (a, b)$, there is a 1–1 correspondence between the set of all number pairs (a, b), whose elements a and b are real numbers and the set $a + ib$, in which a and b belong to C. The reader should have no difficulty in verifying that this correspondence is maintained under addition *and* multiplication, provided we make use of the convention that i^2 is to be replaced by -1. Thus the system $\{C_2$, rational operations$\}$ is isomorphic with the system $\{a + ib$, rational operations$\}$ (compare *48*, 13, 14). Numbers of the form $a + ib$ are also called complex numbers (compare Definition XX); in view of the isomorphism which has just been established this duplication of terminology can not give rise to confusion. The form $a + ib$ is the one in which complex numbers appear in most text-books on algebra.

47. An outlook farther afield. It is a long time since we have had exercise. Before we indulge in it, one further remark has to be made. The reader will recall that at the beginning of *42* (see p. 76) we have deliberately restricted ourselves to plane vectors. The system of complex numbers is adequate for them. But if we want to deal with vectors in a 3-dimensional or 4-, 5- dimensional space, we need still more extended systems. There would be little difficulty in adopting in place of the set C_2 of pairs, the sets C_3, C_4, C_5, etc. of triads, tetrads, pentads, etc. of real numbers (compare *14*). But there would be a good deal of difficulty in finding suitable and adequate definitions for the operations within these sets. And this difficulty increases when we come to deal with the still more general "spaces" to which modern developments in physics have led. The branch of mathematics to which we are here referring has received a great deal of attention. The general field is that of the theory of hypercomplex numbers; in particular, the theory of quaternions has to do with the set C_4.

48. The reality of imaginary numbers.

1. Determine on a line $X'OX$ on which an origin, a direction and a unit are given points corresponding to the rational numbers $+\frac{4}{3}$, $-\frac{5}{7}$, $+\frac{2}{5}$.

2. With the data of the preceding problem determine points corresponding to the irrational numbers $\sqrt{2}$, $\sqrt{\frac{3}{5}}$, $\sqrt{6}$.

3. Construct the product of 2 negative points on the line $X'OX$; also the product of a negative and a positive point.

4. Construct the quotient of a negative point by a positive point; also the quotient of a positive point by a negative point.

5. Prove that the multiplication of points discussed in *41* is commutative.

6. Prove geometrically that the sum of the two pairs (a, b) and $(a, -b)$ is the pair $(2a, 0)$.

7. Show that the sum of several complex numbers represented by the points $P_1, P_2, P_3, \ldots P_k$ can be obtained as follows: with P_1 as initial point, construct a vector equal to OP_2; let its end point be Q_2. With Q_2 as initial point, construct a vector equal to OP_3, leading to an end point Q_3. Continuing in this way, we come to a point Q_k; this point represents the required sum.

8. Prove that the multiplication of number pairs as determined by Definition XIX is associative.

9. Prove that the distributive law holds for the multiplication of number pairs.

10. Restate the definitions for the addition, subtraction, multiplication and division of complex numbers, writing these numbers in the form $a + ib$; show that they are in accordance with rational operations upon algebraic expressions, involving only real numbers, provided it is understood that $i^2 = -1$.

11. Show geometrically that the point P which represents the product of the complex number (a, b) by $(0, 1)$ is obtained by rotating the vector from O to the point $P(a, b)$ through $90°$ about O.

12. Make a list of the first 12 powers of i; interpret the result geometrically.

13. Prove that the set of numbers of the form $a + ib$ in which a and b belong to R, is closed under addition and multiplication.

14. Prove that the 1–1 correspondence between the system C_2 of pairs of real numbers and the system of numbers of the form $a + ib$ is maintained under multiplication and addition (compare p. 86).

49. We return to vectors. From the rule for the multiplication of complex numbers we shall derive a method for multiplying two vectors. This will require that we determine the amount and the direction of the product of two vectors in terms of the amounts and directions of these vectors themselves. It is convenient to associate directly with a complex number the amount and direction of the corresponding vector. This is done as follows:

Definition XXI. The *modulus* of a complex number is the amount of the corresponding vector; the *norm* of a complex num-

ber is the square of the modulus; the *amplitude* (or *argument*) of a complex number is the angle from the positive X-axis to the corresponding vector (see Fig. 17).[1]

The modulus of $a + ib$ is represented by $| a + ib |$; or, if α represents a complex number, by $| \alpha |$ or by $m(\alpha)$; the amplitude

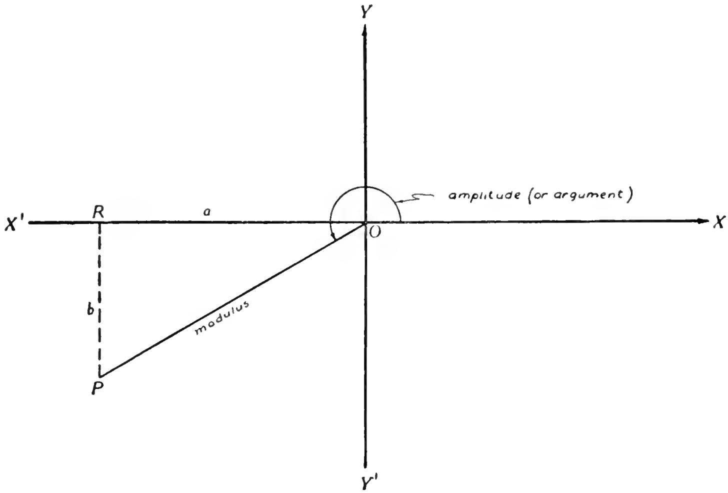


FIG. 17

(argument) of α is written ampl. α (arg. α), or simply $\theta(\alpha)$. The norm of α is abbreviated in the form $N(\alpha)$. We shall also use such notations as $N(a + bi)$, $\theta(a + bi)$, $m(a + bi)$, etc. It is now easy to see that if P represents the complex number $\alpha = a + ib$, then

$$(5.10) \qquad \begin{cases} m(\alpha) = | \alpha | = \sqrt{a^2 + b^2}, \\ N(\alpha) = a^2 + b^2, \\ \theta(\alpha) = \text{ampl. } \alpha = \angle XOP. \end{cases}$$

Our problem consists now in determining a relation between the norms (or the moduli) and amplitudes of two complex numbers and the norm (or modulus) and amplitude of their product.

Take $\alpha_1 = a_1 + ib_1$, and $\alpha_2 = a_2 + ib_2$; let these numbers be

[1] It is important to observe that both the modulus and the norm of a complex number are positive real numbers or zero, and that they are zero only for the complex number (0, 0). From this it follows that not only are the norms of 2 complex numbers equal if their moduli are equal, but conversely the moduli are equal if the norms are.

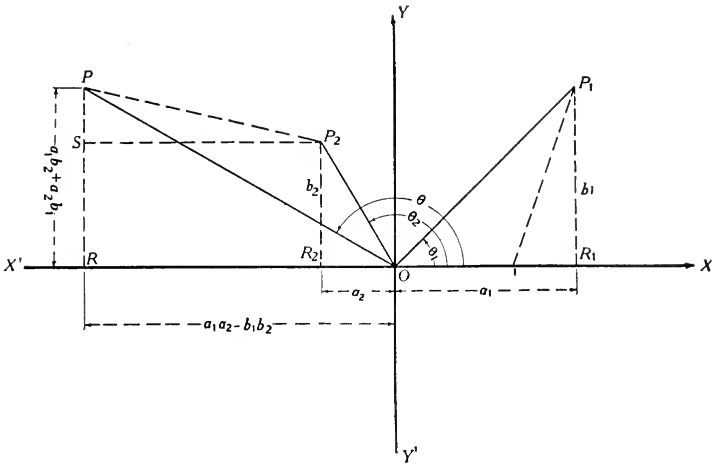


FIG. 18

represented by the points P_1 and P_2 respectively and let their amplitudes be θ_1 and θ_2. And let P represent α, their product, θ the amplitude of the product (see Fig. 18). Then

$$\begin{aligned} N(\alpha) &= (a_1a_2 - b_1b_2)^2 + (a_1b_2 + a_2b_1)^2 \\ &= a_1^2a_2^2 + b_1^2b_2^2 + a_1^2b_2^2 + a_2^2b_1^2 \\ &= (a_1^2 + b_1^2)(a_2^2 + b_2^2) = N(\alpha_1)N(\alpha_2). \end{aligned}$$

From this we conclude in the first place that

(5.11) $$|\alpha_1\alpha_2| = |\alpha_1| \cdot |\alpha_2|.$$

Moreover

$$\begin{aligned} \overline{PP_2}^2 &= \overline{P_2S}^2 + \overline{SP}^2 \\ &= (a_1a_2 - b_1b_2 - a_2)^2 + (a_1b_2 + a_2b_1 - b_2)^2 \\ &= [(a_1 - 1)a_2 - b_1b_2]^2 + [(a_1 - 1)b_2 + a_2b_1]^2 \\ &= [(a_1 - 1)^2 + b_1^2] \cdot [a_2^2 + b_2^2] = \overline{1P_1}^2 \cdot \overline{OP_2}^2. \end{aligned}$$ [1]

From the relations which have been obtained we derive the following:

$$\overline{OP} = \overline{OP_1} \cdot \overline{OP_2} \quad \text{and} \quad \overline{PP_2} = \overline{1P_1} \cdot \overline{OP_2};$$

hence $$\overline{OP} : \overline{OP_2} = \overline{OP_1} : 1 \quad \text{and} \quad \overline{PP_2} : \overline{OP_2} = \overline{1P_1} : 1,$$

or $$\overline{OP} : \overline{OP_1} = \overline{OP_2} : 1 = \overline{PP_2} : \overline{1P_1}.$$

[1] This last calculation need not be carried out in detail; its result is an immediate consequence of what immediately precedes, replacing a_1 throughout by $a_1 - 1$.

This result exhibits the proportionality between the sides of the $\triangle OPP_2$ and those of $\triangle OP_1 1$. It follows that the triangles are similar, and therefore equi-angular; in particular, since $\overline{PP_2}$ and $\overline{1P_1}$ are homologous sides, $\angle P_2OP = \angle 1OP_1 = \theta_1$. Therefore we conclude that $\theta = \angle XOP = \angle XOP_2 + \angle P_2OP = \theta_2 + \theta_1$, i.e. that the amplitude of α is equal to the sum of the amplitudes of α_1 and α_2. We have obtained the following important result:

Theorem X. The modulus (and hence the norm) of the product of two complex numbers is equal to the *product* of their moduli (their norms); the amplitude of the product is equal to the *sum* of the amplitudes:

$$(5.12) \qquad \begin{cases} m(\alpha_1\alpha_2) = m(\alpha_1)m(\alpha_2); \\ N(\alpha_1\alpha_2) = N(\alpha_1)N(\alpha_2); \\ \theta(\alpha_1\alpha_2) = \theta(\alpha_1) + \theta(\alpha_2).[1] \end{cases}$$

Reinterpreting this result in terms of vectors, we find for the product of two plane vectors a new vector whose amount is equal to the product of the amounts and whose direction is obtained by turning one of them through the angle given by the direction of the other.

Since the theorems concerning the norms and those concerning the modulus are equivalent (both modulus and norm being positive and real), only one of the two forms will hereafter be stated.

50. Preparation for an expedition. It is a corollary to Theorem X that the modulus of the square of α is equal to the square of the modulus of α, while the amplitude of α^2 equals 2 times the amplitude of α.

The corollary is expressed in the formulas:

$$m(\alpha^2) = [m(\alpha)]^2, \ \theta(\alpha^2) = 2\theta(\alpha).$$

It follows from this corollary, combined with Theorem X that

$$m(\alpha^3) = m(\alpha^2\alpha) = m(\alpha^2)m(\alpha) = [m(\alpha)]^3$$

and $$\theta(\alpha^3) = \theta(\alpha^2\alpha) = \theta(\alpha^2) + \theta(\alpha) = 3\theta(\alpha).$$

Similarly, if we suppose that

$$m(\alpha^k) = [m(\alpha)]^k \quad \text{and} \quad \theta(\alpha^k) = k\theta(\alpha),$$

we would conclude by using the same method that

$$m(\alpha^{k+1}) = [m(\alpha)]^{k+1} \quad \text{and} \quad \theta(\alpha^{k+1}) = (k+1)\theta(\alpha).$$

Since now our supposition is indeed fulfilled for $k = 3$, it will also

[1] See footnote on p. 92.

hold for $k = 4$; but then, being true for $k = 4$, it holds for $k = 5$, and so forth. We conclude that it holds for every integer k.[1]

In this way we obtain from Theorem X the further conclusion contained in

Theorem XI. The modulus of any positive integral power of a complex number is equal to the same *power* of the modulus of this complex number; the amplitude is equal to the corresponding *multiple* of its amplitude; i.e.

$$(5.13) \qquad m(\alpha^n) = [m(\alpha)]^n \quad \text{and} \quad \theta(\alpha^n) = n\theta(\alpha).$$

51. We reach a peak. We are now in a position to prove a theorem which brings out the important bearing of the extension from the set of real numbers to the set of complex numbers with which we have been concerned in this chapter, upon the question which occupied us in the preceding chapter (see p. 65).

Theorem XII. If α is any complex number and n a natural number, there always exists a complex number X, such that $X^n = \alpha$.

Proof. The number X will be determined if we know $m(X)$ and $\theta(X)$. It follows at once from Theorem XI that these quantities must satisfy the following conditions:

$$(5.14) \qquad [m(X)]^n = m(\alpha),$$

$$(5.141) \qquad n\theta(X) = \theta(\alpha).$$

But $m(\alpha)$ is a positive real number; hence in virtue of Theorem IX (see p. 69) there always exists a single positive real number which satisfies equation (5.14); and it follows from (5.141) that $\theta(X) = \frac{1}{n} \cdot \theta(\alpha)$. Hence both $m(X)$ and $\theta(X)$ exist, and consequently the complex number X exists as asserted in the theorem.

Theorem XII gives us the complete answer to question 1 of page 50 by asserting that the system of complex numbers is closed with respect to the first inverse of the operation of involution.

One important remark remains to be made. In Definition XXI, the number $\theta(\alpha)$ was defined for the complex number α as "the angle from the positive X-axis to the corresponding vector." Is it clear from this statement just what angle $\theta(\alpha)$ is? Clearly we can let a line rotate from the positive X-axis about O until it coincides with OP. But this line can rotate in a clockwise sense, or

[1] We have used here the famous principle of *mathematical induction*, compare Poincaré, *La Science et l'Hypothèse*, pp. 20–28, 112.

in a counter-clockwise sense. (We call the first a *negative* rotation, and the angle so obtained a *negative* angle, and the second a *positive* rotation leading to a *positive* angle; for the case illustrated in Fig. 19 these angles are approximately $+ 220°$ and $- 140°$.) And moreover why should the line stop rotating as soon as it reaches OP? There is no reason why it should. The "angle from the positive X-axis to the corresponding vector" may therefore be any

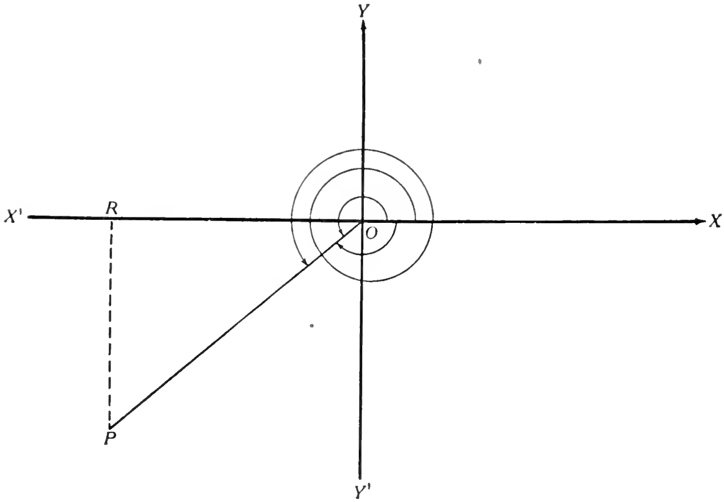


FIG. 19

angle from a denumerable set of positive and negative angles. If θ is the measure in degrees of any angle of that set all the others have measures of $\theta + 360°$, $\theta + 2 \cdot 360°$, $\theta + 3 \cdot 360°$, . . . etc., or $\theta - 360°$, $\theta - 2 \cdot 360°$, $\theta - 3 \cdot 360°$, etc.; they are all contained in the expression $\theta + k \cdot 360°$, where k is an integer, positive, negative or zero. In the theorems on multiplication and involution of complex numbers it has not been necessary to consider these different angles, because the amplitudes of the products and powers that would be obtained by their use would again differ among each other by multiples of $360°$, and therefore give rise to the same complex numbers.[1] For example, (5.13) would tell us that

[1] The critical reader will observe, in connection with Theorem X, that if this theorem is to hold in all cases we may have to use for $\theta(\alpha_1\alpha_2)$ one of the angles to which attention is called here.

$$\theta(\alpha^3) = 3 \cdot [\theta(\alpha) + k \cdot 360°] = 3 \cdot \theta(\alpha) + 3k \cdot 360°;$$

the additional term $3k \cdot 360°$ is of no interest.

This is not the case however with (5.141); it would give, for example, for $n = 3$, $3\theta(X) = \theta(\alpha) + k \cdot 360°$, and hence $\theta(X) = \frac{1}{3} \cdot \theta(\alpha) + k \cdot 120°$. For successive values of k, we obtain therefore angles which differ from each other by 120°; and those, at least for $k = 0$, $k = 1$ and $k = 2$, give rise to different complex numbers. It is clear that $k = 3$, $k = 4$ and $k = 5$ will give angles which differ from the first three by 360° and do not lead therefore to different complex numbers. In the general case of (5.141), we would have

$$n\theta(x) = \theta(\alpha) + k \cdot 360° \quad \text{and} \quad \theta(x) = \frac{1}{n} \cdot \theta(\alpha) + \frac{k \cdot 360°}{n}.$$

For the n successive values $k = 0, 1, 2, \ldots n - 1$ we obtain angles which differ by $\frac{360°}{n}$ and which lead therefore to n different complex numbers. Every other value of k leads to a complex number identical with one of those obtained here. The discussion shows that we can complete Theorem XII as follows:

Theorem XIII. If α is an arbitrary complex number and n a natural number, there always exist exactly n different complex numbers x, such that $x^n = \alpha$. They are given, in the notation following Definition XXI (see p. 88), by their moduli and amplitudes, in the formulas:

$$m(x) = \sqrt[n]{m(\alpha)},$$

$$\theta(x) = \frac{1}{n} \cdot \theta(\alpha) + \frac{k \cdot 360°}{n}, \text{ for } k = 0, 1, 2, \ldots n - 1.$$

Examples

1. Suppose $\alpha = i$ and $n = 4$. It is readily seen that $m(\alpha) = 1$, $\theta(\alpha) = 90° + k \cdot 360°$. Hence $m(x) = \sqrt[4]{1}$, and $\theta(x) = \frac{90°}{4} + k \cdot 90°$, $k = 0, 1, 2, 3$. By using for k integers greater than 3 or less than 0, we obtain for $\theta(x)$ values which differ by multiples of 360° from the values indicated here. There are therefore 4 solutions of the equation $x^4 = i$, viz. the complex numbers corresponding to the points P_1, P_2, P_3 and P_4 in Fig. 20. All these numbers have the modulus 1 and their representative points lie therefore on the circle about the

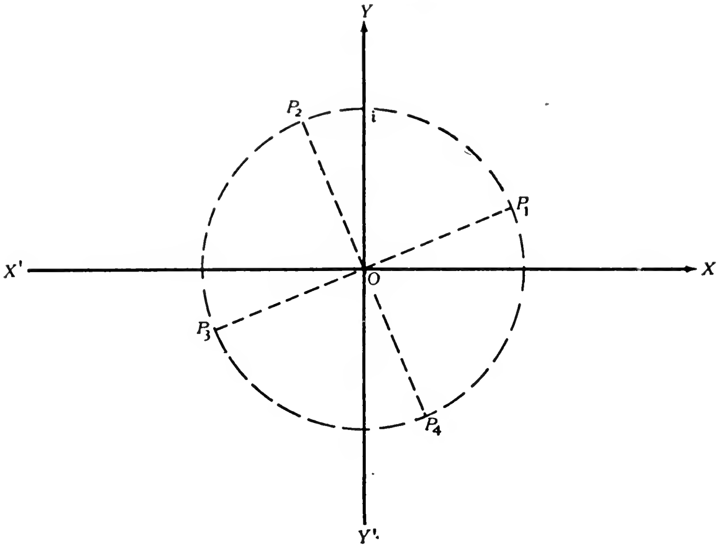


FIG. 20

origin with unit radius; their amplitudes are $\theta_1 = \dfrac{90^\circ}{4} = 22^\circ\,30'$, $\theta_2 = 112^\circ\,30'$, $\theta_3 = 202^\circ\,30'$, $\theta_4 = 292^\circ\,30'$ (see Fig. 20).

2. Suppose $\alpha = 1 + i\sqrt{3}$ and $n = 5$. It follows (see Fig. 21) that $N(\alpha) = 1 + 3 = 4$, and hence that $m(\alpha) = 2$. To determine $\theta(\alpha) = \angle XOP + k \cdot 360^\circ$, we notice that if PQ is drawn in such a way that $\angle QPR = \angle OPR$, then $\triangle OPR$ and $\triangle QPR$ are congruent, so that $PQ = 2$ and $RQ = 1$. Hence the three sides of $\triangle OPQ$ are all equal to 2, the triangle is equilateral and each of its angles is equal to 60°. Therefore $\theta(\alpha) = 60^\circ + k \cdot 360^\circ$. By application of Theorem XIII, we find now:

$$m(x) = \sqrt[5]{2}, \quad \text{and} \quad \theta(x) = \frac{60^\circ}{5} + \frac{k \cdot 360^\circ}{5},\ k = 0, 1, 2, 3, 4.$$

There are now 5 values of x, for each of which $x^5 = 1 + i\sqrt{3}$. They all have a modulus equal to $\sqrt[5]{2}$; their representative points lie on a circle about the origin of radius $\sqrt[5]{2}$. Their amplitudes are

$$\theta_1 = 12^\circ,\ \theta_2 = 84^\circ,\ \theta_3 = 156^\circ,\ \theta_4 = 228^\circ,\ \theta_5 = 300^\circ.$$

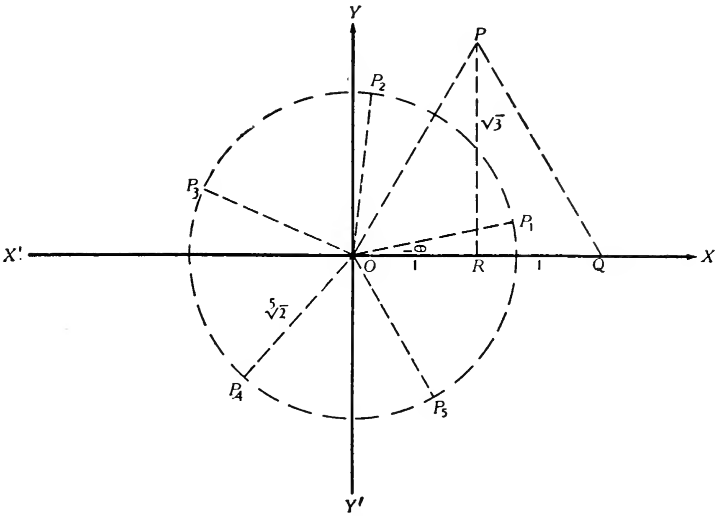


FIG. 21

3. Suppose $\alpha = 1$, and $n = 3$. In this case $m(\alpha) = 1$ and $\theta(\alpha) = 0° + k \cdot 360°$. We find 3 solutions for the equation $x^3 = 1$, each of modulus 1; their amplitudes are 0°, 120° and 240°. Their representative points lie on the unit circle (see Fig. 22). To P_1 corresponds the complex number 1. We can determine the complex numbers corresponding to P_2 and P_3 by means of the discussion used in Example 2. For $\angle ROP_2 = 60°$ and $\angle ROP_3 = 60°$. Leaving the details to the reader, we find that $RP_2 = \frac{1}{2}\sqrt{3}$, $OR = -\frac{1}{2}$ and $RP_3 = -\frac{1}{2}\sqrt{3}$. We conclude that the three solutions of $x^3 = 1$ are $1, -\frac{1}{2} + \frac{1}{2}\sqrt{3}i, -\frac{1}{2} - \frac{1}{2}\sqrt{3}i$.

52. Enjoying the views.

1. Determine the modulus, the norm and the amplitude for each of the following numbers: $i, -1, -i, 1 + i, -1 + i, -1 - i, 1 - i$.
2. Show that the statement in *48*, 11 is an immediate consequence of Theorem X.
3. Develop a construction for the quotient of two plane vectors.
4. Prove that the modulus of the quotient of two complex numbers is the quotient of their moduli, that the norm is the quotient of their norms, and that the amplitude is obtained by subtracting the amplitude of the divisor from the amplitude of the dividend.

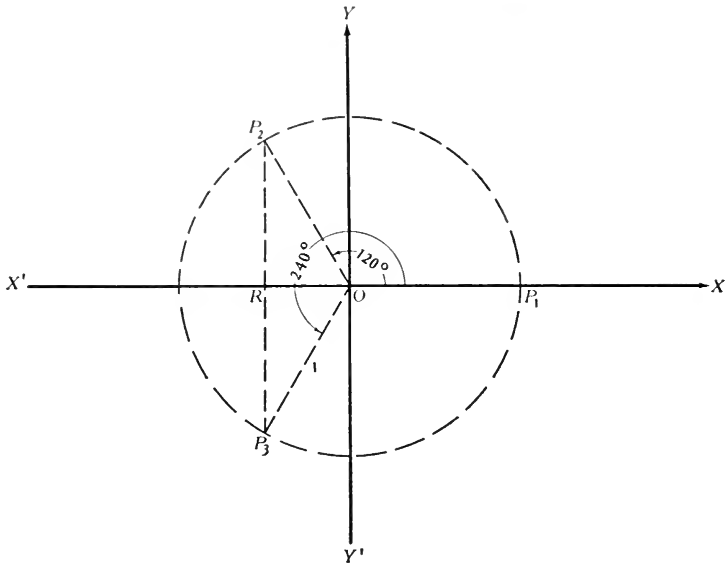


FIG. 22

5. Show that the two complex numbers $a + bi$ and $a - bi$ have equal moduli and equal norms, and that their amplitudes can be so taken as to differ in sign only.

Note. The pair of complex numbers $a + bi$ and $a - bi$ is called a pair of *conjugate complex numbers.*

6. Show that the reciprocal of a complex number whose modulus is 1 is equal to its conjugate.

7. Show geometrically that the sum and product of two conjugate complex numbers are positive real numbers, and that their difference is a normal number.

8. Prove that the conjugates of the sum, difference, product and quotient of two complex numbers are equal respectively to the sum, difference, product and quotient of their conjugates.

9. Determine the 10th power of $1 + i$ (use the result of 1).

10. Determine the moduli and amplitudes of all the complex numbers which satisfy the equation $x^5 = -1 + i$.

11. Determine the 12th power of $-\sqrt{3} + i$.

12. Determine the moduli and amplitudes of all the complex numbers which satisfy the equation $x^4 = -1$.

13. Solve the equation $x^6 = 1$.

53. A bypath. The principal purpose of this chapter was accomplished by the acquisition of Theorems XII and XIII. We wanted to show that the first inverse of the process of involution can always be carried out in the system of complex numbers. As in our earlier work we were not much concerned with the technique of the process by which the numbers which solve the problem can actually and conveniently be determined. The examples and exercises served chiefly to make the meaning of the theorems clear by applying them to concrete instances. Indeed in several of them we did not carry the work through to the actual determination of the complex numbers, but we were satisfied with having found their modulus and amplitude. It remains to study a little further the relation between a complex number and its modulus and amplitude. The problem can be put in this form:

Given a complex number α; determine $m(\alpha)$ and $\theta(\alpha)$. And, conversely, given $m(\alpha)$ and $\theta(\alpha)$, determine α.

Quite apart from technical needs, this problem turns out to be of great general interest and importance. It will therefore be worth our while to follow at least for a short distance this path branching off our main trail.

54. A large domain within reach. The first part of the question is answered at once by reference to our starting point in Chapter I; indeed it has already been answered in formula (5.10).

The introduction of a few new words will facilitate our discussion of the rest. The angles $\theta(\alpha)$ are all obtained by the rotation of a line *from* the positive X-axis to some terminal position; this rotation may be positive (counter-clockwise) or negative (clockwise). We shall say that an angle which is so obtained is in *standard position*. The angle XOP in Fig. 23 is in standard position, the angle YOP is not. If, for an angle in standard position, the terminal side (i.e. the side at which the rotation terminates) lies in the 1st quadrant (or in II, III, IV — compare p. 22), we say that the *angle lies in the 1st quadrant* (or in II, III, IV respectively). Thus the angles of 70°, 100°, 212°, 285° lie respectively in I, II, III and IV.

A great deal of study has been made of the relation between the angle θ and the ratios of the three sides of the right triangle ROP formed by the modulus of the complex number, and its elements a and b (compare Figs. 17, 18, 19, 21, 22). For, once this angle is given, the ratios of any two sides of this triangle are determined,

even though the position of P may vary in accordance with $m(\alpha)$. There are six of these ratios; they are called the *trigonometric ratios* of the angle θ. Each of the six trigonometric ratios has received a name of its own; they can be found in any textbook on

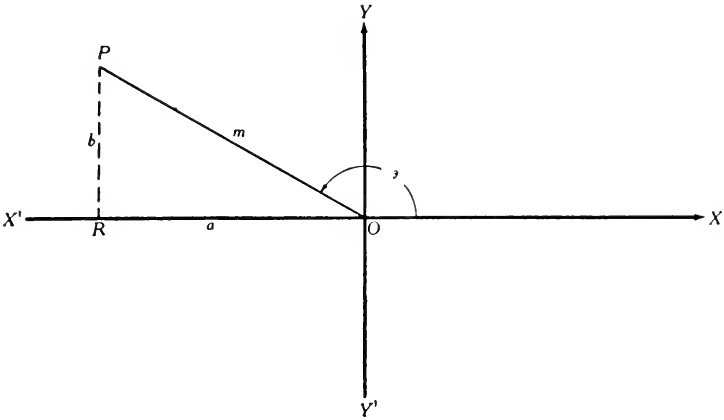


FIG. 23

trigonometry.[1] For the present we shall be concerned only with two of them.

Definition XXII. The ratio of the sides b and m is called the sine of θ (abbreviated sin θ); the ratio of the sides a and m is called the cosine of θ (abbreviated cos θ) (see Fig. 23)

$$(5.15) \qquad \cos\theta = \frac{a}{m}, \qquad \sin\theta = \frac{b}{m}.$$

Since a, b and m are real numbers, m always positive, a and b positive or negative, the sine and cosine of any angle θ are *positive* or *negative* real numbers.

There have been prepared by many authors tabulations of the approximate values of sines and cosines of angles from 0° up to 90°, varying by small amounts. Such a tabulation is called a table of the natural values of the trigonometric functions. A small extract from a "five-place table" is the following:

[1] The origin of the names used to designate the trigonometric ratios is rather interesting, compare e.g. V. Sanford, *A Short History of Mathematics*, p. 298; F. Cajori, *A History of Mathematics*, p. 109.

θ	$\sin\theta$	$\cos\theta$
1°	.01745	.99985
2°	.03490	.99939
3°	.05234	.99863
5°	.08716	.99619
10°	.17365	.98481
20°	.34202	.93969
30°	.50000	.86603
40°	.64279	.76604
45°	.70711	.70711

By means of such tables, our question can be answered, at least approximately. Suppose we had the complex number $\alpha = +4 + 3i$, and we wished to find $m(\alpha)$ and $\theta(\alpha)$. We would find $m(\alpha) = 5$, and hence $\cos\theta(\alpha) = +\frac{4}{5} = +.80000$ and $\sin\theta(\alpha) = +.60000$. Our short table makes it clear that $\theta(\alpha)$ is somewhere between 30° and 40°; more extensive tables would bring the still approximate result $\theta(\alpha) = 36^\circ\ 52'\ 12\frac{1}{2}''$.

If, on the other hand, we had found $\theta(\alpha) = 20^\circ$ and $m(\alpha) = \sqrt{2}$, how could we find α? Since $\frac{a}{m} = \cos\theta$ and $\frac{b}{m} = \sin\theta$, we would find by means of our tables $\frac{a}{\sqrt{2}} = \cos 20^\circ = .93969$, $a = .93969\sqrt{2}$, and $\frac{b}{\sqrt{2}} = \sin 20^\circ = .34202$, $b = .34202\sqrt{2}$.

This is an indication of the technique to be used, but only an indication. For its further development we would have to learn how to use the tables of the natural values, not only for angles from 0° to 90°, but also for positive and negative angles beyond that limited range. Such matters are dealt with in textbooks on trigonometry, of which the woods are full.

As has been seen in Example 3 on page 95 and in some of the problems in *52*, there are some angles whose sine and cosine can be determined without the use of tables. So we find without difficulty:

$$\sin 30^\circ = \frac{1}{2},\quad \cos 30^\circ = \frac{\sqrt{3}}{2};\quad \sin 120^\circ = \frac{\sqrt{3}}{2},\ \cos 120^\circ = -\frac{1}{2};$$

$$\sin 45^\circ = \frac{\sqrt{2}}{2},\ \cos 45^\circ = \frac{\sqrt{2}}{2};\ \sin 135^\circ = \frac{\sqrt{2}}{2},\ \cos 135^\circ = -\frac{\sqrt{2}}{2};$$

$$\sin 60^\circ = \frac{\sqrt{3}}{2}, \cos 60^\circ = \frac{1}{2}; \quad \sin 150^\circ = \frac{1}{2}, \quad \cos 150^\circ = -\frac{\sqrt{3}}{2}.$$

It is also readily seen that

$$\sin 0^\circ = 0, \quad \cos 0^\circ = 1; \quad \sin 90^\circ = 1, \quad \cos 90^\circ = 0;$$
$$\sin 180^\circ = 0, \quad \cos 180^\circ = -1; \quad \sin 270^\circ = -1, \quad \cos 270^\circ = 0.$$

55. We linger a short time. One further digression is unavoidable. Let us consider two complex numbers $\alpha_1 = (a_1, b_1)$ and $\alpha_2 = (a_2, b_2)$ both of which have a modulus equal to 1, i.e. $m(\alpha_1) = m(\alpha_2) = 1$; their representative points will lie on the unit circle. Let us, for simplicity of notation, write θ_1 and θ_2 for $\theta(\alpha_1)$ and $\theta(\alpha_2)$ respectively. Then we shall have, in accordance with (5.15):

$$\cos \theta_1 = a_1, \quad \sin \theta_1 = b_1; \quad \cos \theta_2 = a_2, \quad \sin \theta_2 = b_2.$$

Moreover it follows from Theorem X, that $m(\alpha_1\alpha_2) = 1$, and $\theta(\alpha_1\alpha_2) = \theta_1 + \theta_2$; also from Definition XIX, that

$$\alpha_1\alpha_2 = (a_1, b_1) \cdot (a_2, b_2) = (a_1a_2 - b_1b_2, \; a_1b_2 + a_2b_1).$$

Hence we conclude from Definition XXII, that

$$(5.16) \begin{cases} \cos(\theta_1 + \theta_2) = \cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2, \text{ and} \\ \sin(\theta_1 + \theta_2) = \sin\theta_1 \cos\theta_2 + \cos\theta_1 \sin\theta_2. \end{cases}$$

Theorem XIV. The cosine of the sum of two angles is equal to the product of their cosines diminished by the product of their sines; the sine of the sum of two angles is equal to the sum of the two terms obtainable by taking the product of the sine of either by the cosine of the other.

Formulas (5.16) are known as the *addition formulas* for the sine and cosine. They are the starting point for extensive developments in trigonometry. They enable us, among other things, to extend the usefulness of tables which only run up to 90°; for example,

$$\cos 130^\circ = \cos(90^\circ + 40^\circ) = \cos 90^\circ \cos 40^\circ - \sin 90^\circ \sin 40^\circ$$
$$= -\sin 40^\circ = -.64279.$$

And so we could go on and on and on. But we must remember that we are only making a short digression and that our main road calls us back to its own attractions.

56. We gather strength.

1. Determine the complex number α for which $m(\alpha) = 3$ and $\theta(\alpha) = 40^\circ$.

2. Determine approximate values of the modulus and amplitude of the complex number $2 + i$.

3. It is given that $m(\alpha) = 1$ and $\theta(\alpha) = 20°$; determine α, α^{10}, α^{20}.

4. Determine all the complex numbers which satisfy the equation $x^4 = \alpha$, when α is the number given in 3.

5. Determine geometrically the sine and cosine of 210°, 225°, 240°, 300°, 315°, 330°.

6. Derive from (5.16) formulas for $\cos 2\theta$ and for $\sin 2\theta$ in terms of $\sin \theta$ and $\cos \theta$; verify your result by using $\theta = 30°$ and comparing with the values given at the end of *54*.

7. Prove that it is a consequence of Definition XXII and of (5.10) that for *any* angle θ, the relation

$$(\cos \theta)^2 + (\sin \theta)^2 = 1$$

is valid.

8. Prove that as the angle θ increases from 0° to 90°, $\sin \theta$ increases from 0 to 1, while $\cos \theta$ decreases from 1 to 0.

9. Prove that $\alpha = m(\alpha)[\cos \theta(\alpha) + i \sin \theta(\alpha)]$.

10. Verify Theorem X by means of Theorem XIV.

11. Derive from 6 and 7 the formulas

$$2(\cos \theta)^2 = 1 + \cos 2\theta, \quad 2(\sin \theta)^2 = 1 - \cos 2\theta.$$

12. Calculate the cosine and the sine of 7° 30′ and of 22° 30′.

13. Prove that $(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta$, for any natural number n.

Note. This formula was discovered by the French mathematician Abraham de Moivre (1667–1754); he spent his adult life in England and is known chiefly for his contributions to the theory of probability.[1] The formula is usually referred to as *de Moivre's theorem*.

14. Obtain by means of de Moivre's theorem formulas for $\cos 3\theta$ in terms of $\cos \theta$ and $\sin \theta$.

[1] See F. Cajori, *A History of Mathematics*, p. 245.

CHAPTER VI

FRUITS OF FREEDOM

There are certain primitive ideas which always leave some obscurity in the mind, but which, when their first deductions have once been made, open a field both extensive and free from obstacles. — Lazare Carnot, *Reflexions on the Infinitesimal Analysis*, page 3.

57. A new branch of an old trail. To complete our study of the problem which was stated in Chapter IV, we still have to consider question 2 of page 50, relating to the second inverse of the process of involution. The question as stated there was as follows: If a and b are rational numbers, does there exist a natural number n such that $a^n = b$? It is readily seen that this general question has to be answered in the negative. For while, if we take $a = 10$ and $b = 1000$, we have the obvious answer $n = 3$, on the other hand, if we take $a = 10$ and $b = 1001$, no such natural number n exists. And, as has been observed before, in mathematics one contrary instance is sufficient to destroy a general rule. Instead of being dismayed by such severe insistence upon absolute compliance with the law and the consequent denial of validity to any statement which admits exceptions, even if only a single one, we shall as in our earlier work seek to profit from it. The greater freedom which we have acquired by breaking through the walls of the domain of rational numbers into the wider domains of real numbers and of complex numbers opens up at least a possibility of success. We shall therefore modify the question by transferring it from the domain of rational numbers into that of complex numbers. This requires some preparation; it will be our next concern.

58. Colonization. Let us first undertake to study the following question: If a and b are positive real numbers, does there exist a real number c such that $a^c = b$?

Before we can deal with this new question we have to know what is meant by the symbol a^c when a and c are real numbers. We know what 3^5 stands for, also what is meant by $(-\frac{2}{3})^4$ and by

$(\sqrt[3]{3})^7$. In general, we know what the meaning is of a^n, where a is a real number and n a natural number. But what about 4^{-3}, or $(\frac{7}{2})^{\sqrt{2}}$, or $(\sqrt[4]{3})^{\frac{1}{2}}$? Some of these questions can presumably be answered by means of the results obtained in *35*, 2, 3 and in *39*, 5, 6. The extension of the meaning of a power so as to include powers with negative and zero exponents, required in these exercises, will already have become familiar; the results can be formulated as follows:

Definition XXIII. The power whose exponent is 0 is equal to 1, for every real number, except 0, as base; a power of a real number (except 0) whose exponent is equal to a negative integer is equal to the reciprocal of the power of the same base with the corresponding positive integer as exponent:

$$(6.1) \qquad a^0 = 1, \qquad a^{-n} = \frac{1}{a^n},$$

for every non-zero real number a, and natural number n.[1]

The powers of a number a whose exponents are positive integers form a set isomorphic with the set of powers of a with natural numbers as exponents, with respect to the rational operations.

It follows from this definition that $4^{-3} = \frac{1}{4^3} = \frac{1}{64}$, that $\left(\frac{2}{3}\right)^0 = 1$, etc.

The exercises above referred to have shown that, with these definitions, the laws of exponents of *29* are valid for all integral exponents, positive, zero and negative. We observe in particular that every integral power of 1, whether the exponent be positive, zero or negative, is equal to 1. As a matter of fact, the aim we had before us in framing Definition XXIII was so to interpret powers with negative and zero exponents that the laws of exponents *should* continue to be valid. And in seeking an interpretation for symbols like $5^{\frac{1}{3}}$, $(-4)^{\frac{2}{5}}$, etc. we adhere to the same purpose. It is as if we were sending out from the motherland, the domain of powers with positive integral exponents, an expedition to establish a colony in the strange country of powers with rational exponents. We lay down as a primary consideration that the customs and fundamental

[1] That the value $a = 0$ is exceptional in this definition follows from the discussion in *23* (see pp. 38, 39). The reason for its exclusion from the first part of the definition lies a little deeper.

laws of the motherland shall be adhered to in the organization of social life in the colony.

To begin with, if (4.2) is to hold (see p. 48) we must so define powers whose exponents are rational numbers, that whatever integers p and q are, $a^{\frac{p}{q}} = (a^{\frac{1}{q}})^p$, and in particular so that $a = a^{\frac{q}{q}} = (a^{\frac{1}{q}})^q$; this means that $a^{\frac{1}{q}}$ must be a solution of the equation $x^q = a$. But if a is a positive real number, there exists a single positive real number x which satisfies this equation (see Theorem IX, p. 69). Hence we identify $a^{\frac{1}{q}}$ with this positive real number, and $a^{\frac{p}{q}}$ with its pth power, whether p be a positive integer, a negative integer or zero.

Definition XXIV. The power whose exponent is $\frac{1}{q}$ where q is a positive integer and whose base is the positive real number a is the positive real number which satisfies the equation $x^q = a$; for any integer p the power whose exponent is $\frac{p}{q}$ (reduced to lowest terms) is the pth power of that whose exponent is $\frac{1}{q}$.

In the notation with which we are familiar and which has been used before we have:

$$a^{\frac{1}{q}} = \sqrt[q]{a}; \tag{6.2}$$

$$a^{\frac{p}{q}} = (a^{\frac{1}{q}})^p = (\sqrt[q]{a})^p. \tag{6.21}$$

We shall show now that this definition, if taken as the basis for the organization of the colony, does indeed insure the validity of all the laws of the motherland. It is obvious that $1^{\frac{p}{q}} = 1$ for every rational number $\frac{p}{q}$.

We observe first that an alternative form for $a^{\frac{p}{q}}$ results from these definitions. For if $x^q = a$, and $y = x^p$, then $y = (a^{\frac{1}{q}})^p = a^{\frac{p}{q}}$. On the other hand $y^q = (x^p)^q = (x^q)^p = a^p$, so that $a^{\frac{p}{q}}$ is recognized also as the positive real number which solves the equation $y^q = a^p$, i.e.

$$a^{\frac{p}{q}} = (a^p)^{\frac{1}{q}} = \sqrt[q]{a^p}. \tag{6.22}$$

Let us now formulate explicitly the task which confronts us. We have to show that

1. $$a^{\frac{p}{q}} \cdot a^{\frac{r}{s}} = a^{\frac{p}{q} + \frac{r}{s}},$$

when q and s are positive integers, p and r arbitrary integers.

Suppose then that

(6.3) $$x^q = a \quad \text{and} \quad y^s = a,$$

in other words, that

(6.31) $$x = a^{\frac{1}{q}} \quad \text{and} \quad y = a^{\frac{1}{s}}.$$

From (6.3) follows by (4.2) that

(6.32) $$x^{pqs} = a^{ps} \quad \text{and} \quad y^{rqs} = a^{rq},$$

and hence, by means of (4.1) and (4.4) that

$$(x^p y^r)^{qs} = x^{pqs} y^{rqs} = a^{ps} a^{rq} = a^{ps+rq}.$$

But this means, in virtue of Definition XXIV and (6.22) that

$$x^p y^r = a^{\frac{ps+rq}{qs}} = a^{\frac{p}{q} + \frac{r}{s}}$$

or, by use of (6.31) and (6.21), that

$$a^{\frac{p}{q}} \cdot a^{\frac{r}{s}} = a^{\frac{p}{q} + \frac{r}{s}}.$$

This completes part 1 of our program, for it says that (4.1) holds also for rational exponents. Next we show that

2. $$a^{\frac{p}{q}} \div a^{\frac{r}{s}} = a^{\frac{p}{q} - \frac{r}{s}},$$

without any restriction on the relative magnitude of $\frac{p}{q}$ and $\frac{r}{s}$ but with the same understanding as above, viz. that p and r are arbitrary integers, q and s positive integers. As in 1, we reach (6.32); applying now (4.4) and, instead of (4.1), the formula (4.3) (freed from the restriction $n > m$, by 35, 2), we find

$$\left(\frac{x^p}{y^r}\right)^{qs} = \frac{x^{pqs}}{y^{rqs}} = \frac{a^{ps}}{a^{rq}} = a^{ps-qr}.$$

And now the rest goes as in 1; we find

$$\frac{x^p}{y^r} = a^{\frac{ps-rq}{qs}} = a^{\frac{p}{q} - \frac{r}{s}},$$

or,

$$a^{\frac{p}{q}} \div a^{\frac{r}{s}} = a^{\frac{p}{q} - \frac{r}{s}}.$$

This shows that the colony has maintained (4.3) in force. The next problem is to prove that

3. $$(a^{\frac{p}{q}})^{\frac{r}{s}} = a^{\frac{pr}{qs}}.$$

It follows from Definition XXIV that $a^{\frac{p}{q}}$ is a positive real number; there exists therefore, in virtue of Theorem IX (see p. 69), a positive real number u, such that $u^s = a^{\frac{p}{q}}$. Then, it follows from (6.21) and (6.22) that

(6.4) $$(a^{\frac{p}{q}})^{\frac{r}{s}} = u^r, \text{ and}$$

(6.41) $$(u^s)^q = a^p.$$

By using (4.2) twice, we derive from (6.41) that $u^{sqr} = a^{pr}$, or that $(u^r)^{sq} = a^{pr}$, which means, on the strength of the definition, that $u^r = a^{\frac{pr}{qs}}$. This last statement, in combination with (6.4) shows that indeed

$$(a^{\frac{p}{q}})^{\frac{r}{s}} = a^{\frac{pr}{qs}},$$

or, in other words, that (4.2) is valid in the colony. Finally we have to show that

4. $$a^{\frac{p}{q}} \cdot b^{\frac{p}{q}} = (ab)^{\frac{p}{q}} \quad\text{and}\quad a^{\frac{p}{q}} \div b^{\frac{p}{q}} = \left(\frac{a}{b}\right)^{\frac{p}{q}}.$$

Let x and z be defined by the equations $x^q = a$ and $z^q = b$, so that $x^{pq} = a^p$ and $z^{pq} = b^p$. It follows then from (4.4) that

$$(xz)^{pq} = (ab)^p \quad\text{and}\quad \left(\frac{x}{z}\right)^{pq} = \left(\frac{a}{b}\right)^p.$$

Definition XXIV, in combination with (4.4), leads now to

$$x^p z^p = (ab)^{\frac{p}{q}} \quad\text{and}\quad \frac{x^p}{z^p} = \left(\frac{a}{b}\right)^{\frac{p}{q}},$$

which is equivalent to

$$a^{\frac{p}{q}} b^{\frac{p}{q}} = (ab)^{\frac{p}{q}} \quad\text{and}\quad a^{\frac{p}{q}} \div b^{\frac{p}{q}} = \left(\frac{a}{b}\right)^{\frac{p}{q}};$$

this proves the validity of (4.4) in the new domain.

We have now completed the proof of the statement on page 104, viz. that Definition XXIV does indeed insure the validity of the laws of exponents throughout the domain of powers with rational exponents and positive base.

59. A preliminary clearing of a path. The corresponding problem for irrational exponents is considerably more difficult. We shall limit ourselves to laying down a "suitable" definition for a power of a real number whose exponent is an arbitrary real number, rational or irrational. But we shall not undertake to show that the suitability of this definition consists in enabling us to maintain the laws of exponents which have been shown to be valid for rational exponents.[1] In other words, resuming the parallel of the preceding pages, we shall be content to charter this new colony and to omit the demonstration that this charter will insure the maintenance of the laws of the motherland. But a charter of this kind should not be issued without a good deal of preliminary study. Therefore, returning now to our main highway, we have to clear away some underbrush which obstructs the access to a clear view.

Lemma 1. If a is a real number which is > 1, and q a positive integer, then $a^{\frac{1}{q}} > 1$.

Proof. Let $a^{\frac{1}{q}} = x$; in other words, let $x^q = a$. If $x = 1$, then $a = x^q = 1$; and if $x < 1$, then $a = x^q < 1$. (Compare 3 on p. 28 and 4 in *36*, p. 63.) Both suppositions lead to a contradiction of the hypothesis that $a > 1$. Hence there is nothing left for x, i.e. for $a^{\frac{1}{q}}$, but to be greater than 1; that is what is asserted in the lemma.

Lemma 2. If a is a real number, greater than 1, and r is a positive rational number then $a^r > 1$.

Proof. Let $r = \frac{p}{q}$, where p and q are positive integers; then $a^r = (a^{\frac{1}{q}})^p$, see (6.21). By Lemma 1, $a^{\frac{1}{q}} > 1$; since p is a positive integer, it follows, if we draw once more on 4 in *36*, that $(a^{\frac{1}{q}})^p > 1$.

Lemma 3. If a is a real number, greater than 1, r and s positive rational numbers such that $r > s$, then $a^r > a^s$.

Proof. $a^r - a^s = a^s(a^{r-s} - 1)$, as a consequence of the validity of the laws of exponents (see *58*, 2). Since $r - s$ is a positive rational number, we know from Lemma 2 that $a^{r-s} > 1$, i.e. that $a^{r-s} - 1 > 0$; but s is also a positive rational number, so that $a^s > 1$. From these two facts, we derive, again by means of 4 in *36*, that $a^r - a^s > 0$; this is equivalent to the assertion of the lemma.

[1] See Pierpont, *Theory of Functions of a Real Variable*, vol. 1, pp. 101–108.

Lemma 4. If a_1 is a real number greater than 1 and N any real number, then $a_1^n > N$, provided n is a large enough natural number.

Proof. Since $a_1 > 1$, we can put $a_1 = 1 + e$, where e is positive. Now $(1 + e)^2 = 1 + 2e + e^2 > 1 + 2e$. Hence

$$(6.5) \quad \begin{aligned}(1 + e)^3 &= (1 + e)^2 (1 + e) > (1 + 2e)(1 + e) \\ &= 1 + 3e + 2e^2 > 1 + 3e.\end{aligned}$$

Suppose now that $(1 + e)^k > 1 + ke$. Then

$$\begin{aligned}(1 + e)^{k+1} &= (1 + e)^k(1 + e) > (1 + ke)(1 + e) \\ &= 1 + (k + 1)e + ke^2 > 1 + (k + 1)e.\end{aligned}$$

From (6.5) we can therefore conclude that $(1 + e)^4 > 1 + 4e$; from this we can advance to $(1 + e)^5 > 1 + 5e$. On the strength of this reasoning we conclude that for any natural number n, $(1 + e)^n > 1 + ne$.[1] But to make $1 + ne > N$, we have merely to take $n > \frac{N - 1}{e}$. Therefore if $n > \frac{N - 1}{e}$, we have indeed $a_1^n = (1 + e)^n > 1 + ne > N$.

Suppose, for instance, that $a = 1.001$, and $N = 1000$; then $e = .001$ and $\frac{N - 1}{e} = \frac{999}{.001} = 999{,}000$. We can therefore affirm, without any further calculation, that $(1.001)^{999{,}001} > 1000$, and that $(1.001)^n > 1000$, whenever $n > 999{,}001$.

Lemma 5. If a is a real number greater than 1, and r an arbitrary positive real number, we can make $a^{\frac{1}{n}} < 1 + r$, by taking n large enough.

Proof. Suppose the conclusion were incorrect, i.e. suppose that, no matter how large an integer n were, we should always have $a^{\frac{1}{n}} \geqq 1 + r$. That would mean in virtue of (6.2) that, no matter how large n, we would always have $(1 + r)^n \leqq a$. But this is in direct contradiction with Lemma 4. The hypothesis that Lemma 5 is not valid leads therefore to a contradiction with Lemma 4 of which the validity has already been established. This constitutes a proof of Lemma 5.

The method of proof which was used for Lemma 5 is called the

[1] We have here another example of the principle of "mathematical induction," (compare p. 91). The reader will do well to observe that the same reasoning can be applied to prove that, if $0 < e < 1$, then $(1 - e)^n > 1 - ne$.

"indirect method" or the "reductio ad absurdum." The logical analysis of this method of reasoning is very much worth while; indeed it constitutes one of the most alluring paths which we have encountered on our journey. But it is full of snares and traps. In following it we would get more and more deeply involved in fascinating subterranean difficulties, and delay for a long time our return to daylight. Therefore we had better leave it alone. We shall merely make this simple memorandum: To prove a proposition P by a reductio ad absurdum involves the acceptance of the logical dictum that "either P is true or P is not true," i.e. of the validity and applicability of the "Law of the Excluded Middle." In the Aristotelian logic, this Law of the Excluded Middle is a fundamental canon, applicable to all logical reasoning. It has gone into the warp and woof of all our thinking; it is involved in many of the most important scientific arguments. It is therefore difficult to conceive of its validity being questioned. Nevertheless this has been done in recent years by the *intuitionists* (compare p. 20). It is not surprising that their conclusions are rather startling to the "common sense" of Aristotelian logic. But lest we are led too far along this tempting route, we had better return to our knitting.

It is a simple exercise to prove a series of lemmas, entirely analogous to those which we have just proved, dealing with a real number a less than 1. We shall not state them all explicitly; but for future reference we shall record without proof those which we shall need, viz.:

Lemma 3a. If a is a positive real number less than 1, r and s two positive rational numbers such that $r > s$, then $a^r < a^s$; and

Lemma 5a. If a is a positive real number less than 1, and r an arbitrary positive real number, we can make $a^{\frac{1}{n}} > 1 - r$, by taking n large enough.

As an illustration of Lemma 5, let us take $a = 1000$ and $r = .001$; then we have to show that by taking n large enough we can make $1000^{\frac{1}{n}} < 1.001$. This amounts to saying that we can make $(1.001)^n > 1000$, and we have seen on page 108 that this is indeed the case if $n > 999{,}000$; consequently the 999,001-th root of 1000 is less than 1.001. To illustrate Lemma 5a, let us take $a = .001$ and $r = .001$. We want to show that by taking n large enough we can make $(.001)^{\frac{1}{n}} > .999$, or $(.999)^n < .001$, or again $(\frac{1000}{999})^n > 1000$, which means $(1 + \frac{1}{999})^n > 1000$. Since $1 + \frac{1}{999} > 1.001$,

this condition will surely be fulfilled if $(1.001)^n > 1000$, i.e. if $n > 999{,}000$.

The restriction that r and s in Lemma 3 be positive can now be removed in the following way.

Lemma 3b. If a is a real number greater than 1, r and s negative rational numbers such that $s < r < 0$, then $a^s < a^r$.

Proof. Let $r = -r'$ and $s = -s'$; then $0 < r' < s'$. Moreover, since $a > 1$, $\frac{1}{a} < 1$, so that Lemma 3a is applicable. It tells us that $\left(\frac{1}{a}\right)^{s'} < \left(\frac{1}{a}\right)^{r'}$. Now we are ready to use the law of exponents (it will not be necessary to give specific references — the reader has surely found his way in the colonial domain by now); by means of them the last statement goes over into $a^{-s'} < a^{-r'}$, i.e. $a^r > a^s$. Lemma 3b shows that the conclusion of Lemma 3 remains indeed valid if the condition that r and s be positive is dropped (it is evident on the basis of (6.1) that it surely holds if one of them is zero). A similar procedure shows that this restriction can also be removed from Lemma 3a. We record therefore the following more extensive result:

Theorem XV. If $a > 1$ and real, and $r > s$, both rational, then $a^r > a^s$; if $a < 1$, then $a^r < a^s$.

Corollary. If $a \neq 1$, then $a^r \neq a^s$, unless $r = s$.

The preparations are now complete for the introduction of irrational exponents. Let us then consider the meaning of a^c, where $a > 1$ and c is an irrational number. To say that c is an irrational number is to say that c is a cut in the set R of *rational* numbers without largest element in the lower set (see p. 67). Let this cut be (F, G), a positive or a negative cut; then (1) every $f <$ every g, and (2) elements g and f can be so selected that $g - f$ is less than any desired amount. From (1) follows in conjunction with Theorem XV that $a^f < a^g$. Moreover $a^g - a^f = a^f(a^{g-f} - 1) < a^{g_1}(a^{g-f} - 1)$, where g_1 is any fixed element in G (Theorem XV is used here once more). We conclude now from Lemma 5, that $a^{g-f} - 1$ can be made less than any desired amount r, if we make $g - f < \frac{1}{n}$ and n large enough. But that g and f can be selected so as to accomplish this, is exactly what is asserted in (2); moreover a^{g_1} is fixed in value. Let us now denote by a^F and by a^G the sets of

real numbers obtained when the f in a^f and the g in a^g range over the sets F and G respectively. The sets a^F and a^G have then the following two properties: (1) every element in a^F is less than every element in a^G, (2) elements a^g and a^f can be so selected that $a^g - a^f$ is less than any desired amount. We make now a division of the set R of rational numbers into the two sets A and B, by putting into A every rational number which is less than or equal to any number of the set a^F, and into B every rational number which is equal to or greater than any number of the set a^G. This division constitutes a cut (A, B) in R (compare the argument made on page 68); this cut is the real number which we attach to a^c. In a similar way, we reach an interpretation for a^c, if $0 < a < 1$, but in this case it follows from Theorem XV that the elements of a^F are greater than the elements of a^G; moreover Lemma 5a has to be used instead of Lemma 5. Finally, if $a = 1$, all the elements of a^F and of a^G are equal to 1; the cut (A, B) is then a rational cut and $a^c = 1$. Making the appropriate changes, we can summarize the discussion as follows:

Definition XXV. If the real number c is given by the cut (F, G), and $a > 1$ then the real number a^c is given by the cut (A, B), in which A consists of all rational numbers less than or equal to any element of a^F and B consists of all rational numbers greater than or equal to any element of a^G; if $0 < a < 1$, A consists of the rational numbers less than or equal to any number of a^G, while B consists of the rational numbers greater than or equal to any element of a^F; moreover $1^c = 1$.

Thus we have finally reached a complete definition for the symbol a^c, when a is a positive real number and c an arbitrary real number; this definition is our "charter" of page 107. In our further work we shall accept without proof the validity of the laws of exponents in this extended domain; we shall also assume that the lemmas which have been proved above for rational exponents still hold when the exponents are arbitrary real numbers. Before proceeding to the next stage, it will be useful to consider a few examples and exercises.

Examples. 1. The meaning to be attributed to $5^{\sqrt{2}}$ is the following. Let (F, G) be the cut in R discussed in Example 2 of *34* (see p. 61). We form then the cut (A, B) in which A consists of all rational numbers less than or equal to any number of the form 5^f, and B of all rational numbers equal to or greater than any number of the form 5^g. To A would belong all rational numbers less than

5^1, to B all rational numbers greater than 5^2; to A would belong all rational numbers less than $5^{1.4} = 5^{\frac{7}{5}}$, to B all rational numbers greater than $5^{1.5} = 5^{\frac{3}{2}}$; and so forth. This cut is represented by the symbol $5^{\sqrt{2}}$.

2. The interpretation of $6^{-\sqrt{3}}$ is the same as that of $(\frac{1}{6})^{\sqrt{3}}$, viz. the cut (A, B) in R, in which A consists of all the rational numbers less than or equal to $(\frac{1}{6})^g$, and B of all the rational numbers equal to or greater than $(\frac{1}{6})^f$, where (F, G) is the cut determined by the equation $x^2 = 3$. To A would belong all rational numbers less than the numbers $(\frac{1}{6})^2$, $(\frac{1}{6})^{1.8}$, $(\frac{1}{6})^{1.74}$ etc., to B all rational numbers greater than the numbers $(\frac{1}{6})^1$, $(\frac{1}{6})^{1.7}$, $(\frac{1}{6})^{1.73}$ etc.

60. Surveying the clearing.

1. Prove, by means of (6.1) (see p. 103) and (4.1) to (4.4) (see pp. 48, 49), that $(a^{-n})^{-m} = a^{nm}$, when n and m are positive integers or zero.

2. Determine the real numbers represented by 5^{-2}, $8^{\frac{2}{3}}$, $25^{-\frac{1}{2}}$, $9^{\frac{3}{2}}$, $27^{-\frac{4}{3}}$.

3. Carry out the following operations: $(a^{\frac{1}{3}} - b^{\frac{1}{3}})(a^{\frac{1}{3}} + b^{\frac{1}{3}})$; $(a^{\frac{1}{3}} - b^{\frac{1}{3}})(a^{\frac{2}{3}} + a^{\frac{1}{3}}b^{\frac{1}{3}} + b^{\frac{2}{3}})$; $(a^{-1} - b^{-1})(a^{-2} + b^{-2})$; $(a^{-1} + b^{-1})^2$; $(a^{-1} - b^{-1})^{-2}$. (a and b are positive real numbers.)

4. Determine a natural number n so that $(1\frac{1}{2})^n$ exceeds 500.

5. Determine a natural number n so that $(\frac{9}{10})^n$ is less than .0001.

6. Determine a natural number n so that $\left(\frac{1}{2}\right)^{\frac{1}{n}}$ exceeds .999.

7. Prove that if a is a positive real number, less than 1, and q any positive integer, then $a^{\frac{1}{q}} < 1$.

8. Prove that if a is a positive real number, less than 1, and r a positive rational number, then $a^r < 1$.

9. Prove that if a is a positive real number, less than 1, and b any positive real number, then $a^n < b$, provided n is a sufficiently large natural number.

10. Prove lemmas 3a and 5a (see p. 109).

11. Determine the cut in R which is defined by the symbol $2^{\sqrt{3}}$.

12. Prove, as suggested in the footnote on p. 108, that for any natural number n and any positive real number e less than 1, $(1 - e)^n > 1 - ne$.

61. Another partial conquest. We return now to the problem proposed at the beginning of *58*. Let us begin by taking a special numerical case, and inquire whether there exists a real number x, such that $2^x = 5$. To deal with this question, we use the procedure that served our purpose in *32*. We find that $2^2 = 4$, and $2^3 = 8$;

and we put $a_0 = 2$, $b_0 = 3$. Next we consider the sequence of numbers $2^{2.1}, 2^{2.2}, 2^{2.3}, \ldots 2^{2.8}, 2^{2.9}$. From Theorem XV it follows that this is an ascending sequence of real numbers; there must be two successive ones among them, the first of which is less than 5 while the other exceeds 5. It is a matter of technique to determine which these two are, and we are not now concerned with this aspect of the problem.[1] We would find that $2^{2.3} < 5 < 2^{2.4}$; we put then $a_1 = 2.3$ and $b_1 = 2.4$. Now we consider the numbers $2^{2.31}$, $2^{2.32}, \ldots 2^{2.38}, 2^{2.39}$ and we determine a_2 and b_2 such that $b_2 - a_2 = .01$, and $2^{a_2} < 5 < 2^{b_2}$.

Continuing in this manner, we obtain an ascending sequence of rational numbers $a_0, a_1, a_2, \ldots a_k, \ldots$ and a descending sequence of rational numbers $b_0, b_1, b_2, \ldots b_k, \ldots$, such that all the a's are less than 3, and (1) $b_k - a_k = \dfrac{1}{10^k}$ and (2) $2^{a_k} < 5 < 2^{b_k}$. Moreover $2^{b_k} - 2^{a_k} = 2^{a_k}(2^{b_k - a_k} - 1) < 2^3(2^{b_k - a_k} - 1) = 8(2^{b_k - a_k} - 1)$; it follows from Lemma 5 (see p. 108) that we can make $2^{b_k - a_k} - 1$ less than any desired amount by making $b^k - a^k$ small enough, that is by taking k large enough. It should be clear now to any one who has read *31–34*, how the study of our problem is to be continued. The sequences $a_0, a_1, \ldots a_k, \ldots$ and $b_0, b_1, \ldots b_k, \ldots$ of rational numbers determine a cut (A, B) in R, in which A consists of all rational numbers less than or equal to any number of the first sequence, and B of all rational numbers, equal to or greater than any number of the second. This cut furnishes the answer to our problem. We conclude therefore that there exists a real number x such that $2^x = 5$. The discussion of this special case can be carried over, word for word, to the general case, in which we ask for a real number c such that $a^c = b$, where a and b are positive real numbers and $a \neq 1$.[2] Thus we reach the following result:

Theorem XVI. If a and b are positive real numbers, $a \neq 1$, there exists a single real number c such that $a^c = b$; if $a = 1$, and

[1] It is not difficult to see that $2^{2.1} < 5$, $2^{2.2} < 5$, $2^{2.3} < 5$, and $2^{2.4} > 5$. For these are equivalent to the inequalities $2^{21} < 5^{10}$, $2^{22} < 5^{10}$, $2^{23} < 5^{10}$, $2^{24} > 5^{10}$. But $5 = \dfrac{4 \cdot 5}{4}$, and therefore $5^{10} = 4^{10}(\frac{5}{4})^{10} = 2^{20} \cdot (\frac{5}{4})^{10}$. Therefore it remains to show that $2 < (\frac{5}{4})^{10}$, $4 = 2^2 < (\frac{5}{4})^{10}$, $8 = 2^3 < (\frac{5}{4})^{10}$, $16 = 2^4 > (\frac{5}{4})^{10}$; this can be done easily if we calculate $(\frac{5}{4})^{10} = (1 + \frac{1}{4})^{10}$ by the binomial theorem.

[2] If this argument is carried through, it will be found that Lemma 5a has to be used instead of Lemma 5 if $a < 1$.

$b \neq 1$ there is no real number c for which $a^c = b$, and if $a = b = 1$, $a^c = b$ for every real number c.

That there is but a single number c which solves the problem, in case $a \neq 1$, is seen as follows. Suppose that $a^c = b$, and $a^d = b$, so that $a^d = a^c$. The corollary to Theorem XV would then lead to the conclusion that $d = c$. The last part of the theorem is an immediate consequence of the last part of Definition XXV.

This result has been made possible by the greater freedom which we secured when in Chapter IV we advanced from the restricted domain of rational numbers to the wider one of real numbers; it is an additional justification for the labor we went through to secure this freedom. The quest for the second inverse of involution (see p. 50) has been crowned so far with a partial success: there exists a second inverse within the field of real numbers, provided a and b are *positive* real numbers, $a \neq 1$.

As in much of our earlier work, we have been satisfied to show the existence of the number c which solves the equation $a^c = b$, and we have not concerned ourselves with the technique by means of which it can be determined. This technique involves a good many features which deserve attention quite apart from their technical significance. We can not well afford to miss them because they are of considerable assistance on our further journey. As a traveler through the wilderness is often repaid when he departs from his trail to climb an eminence from which he may gain a view over the surrounding country, so we shall not regret this deviation from our path into a more technical realm.

62. Logarithms. Suppose that we know the two real numbers x and y which solve the equations $a^x = p$ and $a^y = q$, where a, p and q are positive real numbers, $a \neq 1$. These numbers furnish at once the solution of several similar equations, viz. of the equations

$$(6.6) \qquad a^c = pq, \qquad a^c = \frac{p}{q} \qquad \text{and } a^c = p^k,$$

where k is any real number. For, on the one hand it follows from 58 and 59, that

$$(6.7) \qquad a^{x+y} = pq, \qquad a^{x-y} = \frac{p}{q}, \qquad \text{and } a^{kx} = p^k;$$

on the other hand, we know from Theorem XVI that equations (6.6) possess only one solution. We can therefore conclude that

the solutions of these equations are $x + y$, $x - y$ and kx respectively. Of this fact we can make use in a very important way to determine the product and quotient of two real numbers and any power of a real number with real exponent. We must look into this a little further; a technical name will help in this study.

Definition XXVI. If a and b are positive real numbers, $a \neq 1$, and $a^c = b$, then c is called the *logarithm* of b with respect to the base a.[1]

The introduction of the word logarithm makes it possible to restate the sentence "b is equal to the power of a with the exponent c," in which b is the subject in such a way that c becomes the subject; we obtain then the following form: "c is the logarithm of b with respect to the base a." The abbreviated form for this is: $c = \log_a b$. It carries out the rule (man muss immer umkehren, see p. 40) which led us to the problem with which this chapter is concerned. When we write $5 = \log_2 32$, we have inverted the familiar statement $32 = 2^5$.

Using our new toy, we can put the conclusion derived from (6.7) in the following form: If x is the logarithm of p with respect to the base a and y the logarithm of q, then $x + y$ is the logarithm of pq, $x - y$ the logarithm of $\frac{p}{q}$ and kx the logarithm of p^k. We record the result briefly as follows:

Theorem XVII. If a, p and q are positive real numbers, $a \neq 1$, and k an arbitrary real number, then

(1) the logarithm of pq with respect to the base a is equal to the sum of the logarithms of p and q with respect to that base;

(2) the logarithm of $\frac{p}{q}$ with respect to the base a is equal to the logarithm of p diminished by the logarithm of q;

(3) the logarithm of p^k with respect to the base a is equal to k times the logarithm of p.

[1] Concerning the meaning and origin of the word logarithm, the invention and development of logarithms, the reader will find interesting material in the *Napier Tercentenary Memorial Volume*, by E. S. Knott. A brief account is found in Sanford's *A Short History of Mathematics*, pp. 191–197, in Cajori's *History of Mathematics*, p. 161, and in every other book on the history of mathematics. It is worthy of notice that it is only a little over 300 years since logarithms were invented by the Scotchman John Napier. The tercentenary celebration of this invention in Edinburgh in July 1914 was the last international scientific meeting before the outbreak of the World War.

In abbreviated form, this becomes:

$$(6.8)\quad \log_a pq = \log_a p + \log_a q, \quad \log_a \frac{p}{q} = \log_a p - \log_a q,$$
$$\log_a p^k = k \log_a p.$$

To give this Theorem effective significance there have been prepared tables which enable the experienced user to find readily the approximate value of the logarithm of a positive real number with respect to the base 10. There is a great variety of such tables. The most common ones are so-called *five-place* tables; from them can be found, accurately to 5 decimal places, the logarithms to the base 10 of numbers which can be written with at most 5 significant digits, such as 75,823 or 42.065, or .0089321. With the aid of such tables, we find for instance that, accurate to 5 decimal places, $\log_{10} 75{,}823 = 4.87980$, $\log_{10} 42.065 = 1.62392$.

With a little practice, these tables can also be used to solve the inverse problem, viz. that of finding 5 digits of the number whose logarithm with respect to the base 10 is given to at most 5 decimal places. For example, they enable us to find x, if $\log_{10} x = 1.2074$, i.e. if $x = 10^{1.2074}$. A more detailed study of the way in which such tables are used, and of the way in which they are made, is not included in our itinerary. They are very interesting side trips, but they call for extra charges of time and attention. The second one is rather extensive; both are very much worth while, but we must leave them out. For various scientific purposes, very extensive tables are required; there are in existence 14-place, and even 20-place tables, principally for use in astronomical calculations. We will be satisfied with a small extract from a five-place table, which will help us to understand their use. Every one who has read the preceding chapters with some care can construct the following table:

Number	logarithm to the base 10
.001	− 3.00000
.01	− 2.00000
.1	− 1.00000
1	0.00000
10	1.00000
100	2.00000
1000	3.00000

This alone is not yet very useful; let us supplement it with some entries adapted from a 5-place table,[1] and selected to suit our purpose:

n	$\log_{10} n$
12536	4.09816
27649	4.44168
35284	4.54758
44232	4.64574
55450	4.74390
63633	4.80368
76467	4.88336
91358	4.96074

From it we find, with the approximation mentioned above, that $\log_{10} 12536 = 4.09816$ and $\log_{10} 35284 = 4.54758$. By means of Theorem XVII, (1), it follows that $\log_{10} (12536 \times 35284) = 4.09816 + 4.54758 = 8.64574$, and hence that

$$12536 \times 35284 = 10^{8.64574} = ?$$

Our miniature table tells us that $10^{4.64574} = 44232$. Now we make use of the laws of exponents which hold in the colonized domain (see pp. 48, 103, 111), to conclude that

$$\begin{aligned} 12536 \times 35284 = 10^{8.64574} &= 10^{4+4.64574} = 10^4 \times 10^{4.64574} \\ &= 10{,}000 \times 44{,}232 = 442{,}320{,}000, \end{aligned}$$

in which only the first five digits of the result are accurate. To obtain a more accurate result, more extensive tables are required. We must again remember that it is the procedure we are principally interested in, rather than the numerical result. Let us take another example.

From the table, and Theorem XVII, (2), it follows that

$$\begin{aligned} \log_{10}\left(\frac{35{,}284}{55{,}450}\right) = 4.54758 - 4.74390 &= (5.54758 - 4.74390) - 1 \\ &= .80368 - 1. \end{aligned}$$

We have to find now $10^{.80368-1} = 10^{.80368} \div 10^1$. The table informs us that $10^{4.80368} = 63{,}633$; from this information we deduce (see

[1] Very convenient among the smaller tables is the book called the *Macmillan Logarithmic and Trigonometric Tables;* it contains much useful information besides the 5-place table of logarithms.

pp. 48, 103, 111) that $10^{.80368} = 63{,}633 \div 10^4$, and hence that $\frac{35{,}284}{55{,}450} = 63{,}633 \div 10^5 = .63633$, accurate to 5 decimal places.

As a further example, let us compute $27{,}649^2$. From Theorem XVII, (3) and the table we derive the facts that $\log_{10} 27{,}649^2 = 2 \log_{10} 27{,}649 = 2 \times 4.44168 = 8.88336$; therefore $27{,}649^2 = 10^{8.88336} = 10^{4+4.88336} = 10^4 \times 10^{4.88336} = 76{,}467 \times 10^4 = 764{,}670{,}000$, in which only the first five digits are certain.

The examples which have been given so far are only of limited interest because the results obtained in them could be secured without the use of logarithms by any one who has the skill and the patience to do multiplication and division. As a final example we shall therefore take a problem which can not be solved by mere patience and skill in the use of these processes, the bugbears of many people, old and young; let us calculate $\sqrt[5]{63{,}633}$. By means of (6.2), Theorem XVII, (3) and the table, we find

$$\log_{10} \sqrt[5]{63{,}633} = \log_{10} 63{,}633^{\frac{1}{5}} = \tfrac{1}{5} \log_{10} 63{,}633 = \tfrac{1}{5} \times 4.80368$$
$$= .96074.$$

Therefore $\sqrt[5]{63{,}633} = 10^{.96074} = 10^{4.96074-4} = 10^{4.96074} \div 10^4$
$$= 91{,}358 \div 10{,}000 = 9.1358.$$

From this example we see that we have really increased our power in numerical calculation by the study of the second inverse of the process of involution. Jacobi knew what he was talking about, when he said . . .; the reader will know by now what he said and what he meant. Standing on your head is not a foolish thing to do, provided you have something in your head (see p. 41).

The following exercises will give the reader an opportunity to try his hand at the new game and to become acquainted with the useful and fascinating tool which the invention of logarithms has put at the disposal of mankind. Is it not a pity that so many people have to go through life without having enjoyed the unique pleasure which it affords? It is hoped the reader will agree that it justifies our short excursion into the somewhat more technical field.

63. Taking part in the game.

1. Use the miniature table of p. 117 to determine 44232×12536 (compare your result with the answer a younger member of your family

gets at the expenditure of a long and laborious use of pencil and a lot of paper).

2. By use of the table and the additional information that $\log_{10} 12450 = 4.09516$, determine 35284^2. (Another chance to assist in the education of a young friend!)

3. On the supposition that $\log_{10} 23014 = 4.36200$, calculate $\frac{63633}{27649}$.

4. How do you account for the fact that all the logarithms in the miniature table begin with 4?

5. Prove that $\log_a pqr = \log_a p + \log_a q + \log_a r$.

6. Being given that $\log_{10} 44{,}632 = 4.64965$, determine $\sqrt[7]{35{,}284}$.

7. Prove that $\log_a \sqrt[k]{\frac{p^2 q}{r}} = \frac{2 \log_a p + \log_a q - \log_a r}{k}$.

8. Determine $\log_2 4$, $\log_2 8$, $\log_2 16$, $\log_2 2$, $\log_2 1$, $\log_2 \frac{1}{2}$, $\log_2 \frac{1}{4}$.

9. Prove that, for every positive real number a, $\log_a a = 1$, $\log_a 1 = 0$.

10. What is the meaning of $10^{\log_{10} 7}$? (In preparation for the consideration of this simple question of formidable appearance, the reader might reflect on the following profound problems: What is the name of the man whose name is Jones? how long did the 80-years war last?)

11. Determine the relation which holds between the logarithms of numbers to the base 2 and their logarithms to the base $\frac{1}{2}$.

12. Explain why 1 can not be used as the base of a system of logarithms.

13. Evaluate $5^{\log_5 8}$ without doing any calculating. Set up a general statement of which your answers to this question and to 10 above are special cases.

14. Set up and prove a general proposition of which your answer to 11 above is a special case.

64. We reach higher altitudes. The reader will perhaps have got some idea, particularly if he has stopped for the exercises, why 10 is an especially useful number to use as the base of logarithms. Theoretically any fixed base except 1 is as good as any other one (certainly a base has to be fixed if anything is to be built on it; and what is the use of a base on which nothing can be built?). In actual practice, calculations are much simplified for users of the decimal notation, and the usefulness of short tables is greatly enhanced, if 10 is used as a base. But it was not the base used by Napier, the inventor of logarithms, in his famous work entitled *Mirifici logarithmorum canonis descriptio.* Logarithms with respect to the base 10, usually called *Common logarithms*, were first calculated by *Henry Briggs* in England, a contemporary of Napier,

who published in 1624 the *Arithmetica logarithmica*, giving 14-place logarithms of the numbers from 1 to 20,000, and from 90,000 to 100,000; the gap was filled in 1628 by Adrian Vlacq in Holland.[1]

What base did Napier use, and why did he not use 10? The answer to the second of these questions will involve us in the side trip, which was declined on page 116; we must leave it unanswered. The first one can readily be answered: He used the base e.[2] But what is e? It is a positive real number that can be calculated to as high a degree of accuracy as may be desired by taking enough of the denumerable infinitude of terms of the series

$$(6.9) \quad 1 + \frac{1}{1} + \frac{1}{1.2} + \frac{1}{1.2.3} + \frac{1}{1.2.3.4} + \frac{1}{1.2.3.4.5} + \ldots,$$

which is called the exponential series. The sum of the first three terms is 2.5; accurately to 8 decimal places e is equal to 2.71828183. It is another one of the real numbers which like π (see p. 67) is not contained among the numbers which are obtained from the first inversion of involution. It can be defined as a cut in the system R of rational numbers; but this definition is not very accessible for us just now. We shall have to be content with the series (6.9) as a definition of the number e; and we shall state, without proof, one of its remarkable properties. Logarithms with respect to the base e are called *natural logarithms*, also *Napierian logarithms*.

At the beginning of 58 the question was asked as to the meaning to be attributed to a^c, when a and c are both real numbers. After a good deal of preliminary work this question was answered, for arbitrary real c and positive real a, by Definitions XXIV and XXV (see pp. 104, 111), and for integral c and arbitrary real a by Definition XXIII (see p. 103). What we really want and should have, once we have reached the domain of complex numbers, is a suitable definition of a^c, when a and c are arbitrary complex numbers. This is made possible by the property of the number e that we are now going to state, viz. by the formula

$$(6.01) \qquad e^{iy} = \cos y + i \sin y.$$

This formula, discovered by the famous Swiss mathematician *Leonhard Euler* (1707–1783),[3] is one of the gems of elementary

[1] See Cajori, *A History of Mathematics*, p. 164.

[2] This statement is only a rough approximation of the actual fact; compare Sanford, *op. cit.*, pp. 193–197.

[3] Euler is one of the outstanding figures in the history of mathematics. His work

mathematics; unfortunately we are not within reach of its proof at the present stage of our journey. The symbol i which occurs in it is the normal unit, i.e. the complex number (0, 1) with which we are familiar from Chapter V (see in particular *46*). We shall understand y to represent an arbitrary real number, so that cos y and sin y are the trigonometric functions with which we also became acquainted in Chapter V (see Definition XXII, p. 98).

There must however be one further understanding. The formula would tell us, for instance, that $e^{2i} = \cos 2 + i \sin 2$. But what are cos 2 and sin 2 to mean? Are they to be the cosine and sine of an angle of 2°? No, they are not; they are to be the cosine and sine of an angle of 2 *radians;* and the *radian* is the angle which, constructed with the center of any circle as vertex, intercepts on the circumference of this circle an arc equal to the radius of the circle. We have to understand in formula (6.01) that y is the measurement of an angle *in radians.* Every one knows that if the radius of a circle is R, then its circumference is $2\pi R$, i.e. that the ratio of circumference to radius is equal to 2π. Consequently the ratio of an angle of 360° to one radian is also 2π; in other words $360° = 2\pi$ radians. From this we find that $180° = \pi$ radians, $90° = \frac{\pi}{2}$ radians, $60° = \frac{\pi}{3}$ radians, $1° = \frac{\pi}{180}$ radians, etc. Changing the measurement of an angle from degrees into radians, or vice versa, is analogous to changing temperature measurements from centigrade to Fahrenheit, or length measures from inches to centimeters, or money values from dollars to francs, or marks — only it is much easier than any of these things; it is a good simple game. Euler's formula remains valid if y is an arbitrary complex number; its interpretation in that case requires however an extension of the definition of trigonometric functions to complex numbers — another extra trip which has to be omitted.

Euler's formula (6.01) with the above interpretation of the symbols which occur in it enables us to go a good deal further afield in the domain of complex numbers, provided we can be assured that the laws of exponents, our right arm in earlier colonizing

affected profoundly practically every part of the subject. His collected writings have been in course of publication by the Swiss Society of Naturalists since 1911. So far 22 octavo volumes have been published; it is estimated that the complete set will contain over 80 volumes.

The formula, here attributed to him, apparently was found in essence at an earlier date by Roger Cotes (1682–1716); compare Sanford, *op. cit.*, p. 190.

expeditions, still reign supreme; for in that case we can say

(6.02) $$e^{x+iy} = e^x \cdot e^{iy} = e^x(\cos y + i \sin y),$$

when x and y are arbitrary real numbers; we have then an interpretation for a^c, when c is an arbitrary complex number, provided $a = e$. We shall proceed on the assumption that the domain C_2 does recognize these fundamental laws.

Suppose then that for the arbitrary non-zero complex number α, we have $m(\alpha) = r$ and $\theta(\alpha) = y$ (compare pp. 88, 101); then

$$\alpha = r(\cos y + i \sin y) = \frac{r}{e^x} \cdot e^x(\cos y + i \sin y) = \frac{re^{x+iy}}{e^x}.$$

Hence $\alpha = e^{x+iy}$, provided $r = e^x$. Now we use our new toy; it enables us to transform the last statement into the equivalent form:

$$\log_e \alpha = x + iy, \text{ provided } x = \log_e r; \text{ i.e.,}$$
$$\log_e \alpha = \log_e r + iy = \log_e m(\alpha) + i\theta(\alpha).$$

This leads us to the following definition:

Definition XXVII. The natural logarithm (i.e. the logarithm with respect to the base e) of the non-zero complex number α is any complex number whose real part is the logarithm of the modulus of α and whose normal part is i times any value of the amplitude of α.[1]

We recall (see p. 92) that $\theta(\alpha)$ has a denumerable infinitude of values; consequently $\log_e \alpha$ has a denumerable infinitude of values.

This new definition demonstrates the manner in which the introduction of complex numbers has enlarged our freedom, for we are now able to determine the natural logarithm of *any* complex number. Let us consider a few examples.

To find log (-1), we remember (see *52*, 1) that $m(-1) = 1$, and $\theta(-1) = 180° + k \cdot 360° = \pi + 2k\pi$ radians; also that $\log_e 1 = 0$ (see *63*, 9). We obtain then

$$\log_e (-1) = 0 + (\pi + 2k\pi)i = \pi i + 2k\pi i;$$

written in exponential form it becomes $e^{\pi i} = e^{3\pi i} = e^{5\pi i} = \ldots = -1$, a remarkable relation which is an immediate consequence

[1] Since r, the modulus of α, is a positive real number (see the footnote on p. 88) and e is also a positive real number, the meaning of log r is clear from Definition XXVI (see p. 115).

of Euler's formula (6.01). It follows furthermore that $e^{2\pi i} = (e^{\pi i})^2 = (-1)^2 = 1$; and, more generally, that $e^{2n\pi i} = (e^{2\pi i})^n = 1^n = 1$, for any integral value of n. We record this last result for future use.

(6.011) $\qquad e^{2n\pi i} = 1$, for every integer n.

To find $\log_e(+1 + i\sqrt{3})$, we recall (see Example 2, p. 94) that $m(+1 + i\sqrt{3}) = 2$, and $\theta(+1 + i\sqrt{3}) = 60° + k \cdot 360° = \dfrac{\pi}{3} + 2k\pi$; we obtain at once

$$\log_e(+1 + i\sqrt{3}) = \log_e 2 + \frac{\pi i}{3} + 2k\pi i, \text{ where } k \text{ is any integer.}$$

65. The final dash. We can now readily take the final step, necessary to reach the goal which we have set ourselves in this chapter, viz. that of solving for c the equation $\alpha^c = \beta$ in which α and β are arbitrary complex numbers. Let us begin by inquiring whether Definition XXVII is of such character as to maintain the validity of Theorem XVII, when we are dealing with the natural logarithms of arbitrary complex numbers. The fact that these logarithms have a denumerable infinitude of values introduces a complication, different from anything we have met so far; moreover we still have to devise a definition for α^c, when both α and c are complex numbers, which will maintain the laws of exponents intact. We shall suppose now that all the number symbols we are using represent complex numbers, unless the contrary is stated; also that all the logarithms which are used are natural logarithms (i.e. with respect to the base e), unless otherwise specified. We have to inquire next what we can conclude from the relation $pq = r$.

We take first an example, viz. the relation $-i(-1-i) = -1 + i$. We know, at any rate we should be able to find without difficulty, that

$$m(-i) = 1, \text{ and } \theta(-i) = 270° + n \cdot 360° = \frac{3\pi}{2} + 2n\pi;$$

also that

$$m(-1-i) = \sqrt{2}, \text{ and } \theta(-1-i) = 225° + h \cdot 360° = \frac{5\pi}{4} + 2h\pi,$$

and

$$m(-1+i) = \sqrt{2}, \ \theta(-1+i) = 135° + l \cdot 360° = \frac{3\pi}{4} + 2l\pi.$$

Consequently,

$$\log(-\ i) = \frac{3\pi i}{2} + 2n\pi i;\ \log(-\ 1 - i) = \log\sqrt{2} + \frac{5\pi i}{4} + 2h\pi i;$$

$$\log(-\ 1 + i) = \log\sqrt{2} + \frac{3\pi i}{4} + 2l\pi i.$$

In these statements n, h and l are all three arbitrary integers, positive, zero or negative (they are exceptions to the agreement that all the number symbols are to represent arbitrary complex numbers).

There is no difficulty with respect to the relation between the real parts, for the real parts of $\log(-\ i)$ and $\log(-\ 1 - i)$ are equal to 0 and $\log\sqrt{2}$ respectively; and their sum is $\log\sqrt{2}$, which is the real part of $\log(-\ 1 + i)$. With respect to the normal parts, we observe that $\frac{3\pi i}{2} + \frac{5\pi i}{4} = \frac{11\pi i}{4} = \frac{3\pi i}{4} + 2\pi i$; i.e. if we take $n = h = 0$, and $l = 1$, then the sum of the normal parts of the logarithms of $-\ i$ and $-\ 1 - i$ is equal to the normal part of the logarithm of their product, as it should be if Theorem XVII, (1) is to hold. Similarly if we take $n = 0$, $h = 1$ (or $n = 1$, $h = 0$), then $l = 2$ will do. And in general, if we put $l = n + h + 1$, the desired relation holds between the normal parts of the logarithms under consideration.

Let us now look at the general case $pq = r$. We know (see Theorem X, p. 90), that

$$m(p) \cdot m(q) = m(r), \text{ and} \tag{6.021}$$

$$\theta(p) + \theta(q) = \theta(r). \tag{6.022}$$

Here $\theta(p)$, $\theta(q)$ and $\theta(r)$ each have a denumerable infinitude of values; and we are to understand the relation in this way that, no matter what values are chosen for $\theta(p)$ and $\theta(q)$, there is always a corresponding value for $\theta(r)$, such that $\theta(p) + \theta(q) = \theta(r)$. This being clearly understood, we can proceed. For, $\log p = \log m(p) + i\ \theta(p)$, $\log q = \log m(q) + i\ \theta(q)$, $\log r = \log m(r) + i\ \theta(r)$; and it follows from (6.021) and Theorem XVII, (1), since $m(p)$, $m(q)$ and $m(r)$ are positive real numbers (see p. 88) that $\log m(p) + \log m(q) = \log m(r)$. In combination with (6.022) and Definition XXVII this enables us to affirm that

$$\log p + \log q = \log r,$$

in the sense that, no matter which of the denumerable infinitude of values of $\log p$ and $\log q$ are selected, there is always a value of $\log r$, such that this relation holds, i.e. so that Theorem XVII, (1) is valid; but it need not hold if $\log p$, $\log q$ and $\log r$ are all three selected arbitrarily from among the infinitudes of values they have.

An entirely analogous conclusion holds for Theorem XVII, (2); its justification is left to the reader.

As to part 3 of the theorem, we have first to define the meaning of p^k, when p and k are complex numbers. For this purpose we begin by observing that $p = e^{\log p}$ (remember the man whose name is Johnson, see *63*, 10). If then the laws of exponents are to remain in force, we must have $p^k = (e^{\log p})^k = e^{k \log p}$; this we take as the definition of p^k in the field of complex numbers.

Definition XXVIII. If p and k are arbitrary complex numbers, the symbol p^k designates any one of the denumerable infinitude of values represented by $e^{k \log p}$.

An equivalent form of this definition is the following:

Definition XXVIIIa. If p and k are arbitrary complex numbers, the symbol $\log p^k$ designates any one of the infinitude of values represented by $k \log p$.

The validity of part 3 of Theorem XVII is an immediate consequence of these definitions.

An example will help us to understand the significance of this interpretation of the symbol p^k. Let us consider the expression $(-1-i)^{1+i}$. The original definition of a power with a natural number as exponent (compare *28*, p. 47) is quite useless here; and so are the extensions which were introduced in the present chapter (see *58*, pp. 103–107). We must have recourse to the most recent (and final) extension introduced in Definition XXVIII. In accordance with it, we find that

$$(6.03) \qquad (-1-i)^{1+i} = e^{(1+i)\log(-1-i)}$$

On page 124 we found that

$$\log(-1-i) = \log\sqrt{2} + i\left(\frac{5\pi}{4} + 2h\pi\right).$$

Hence the exponent on the right side of (6.03), i.e. $\log(-1-i)^{1+i}$, is equal to $(1+i)\left[\frac{1}{2}\log 2 + i\left(\frac{5\pi}{4} + 2h\pi\right)\right]$

$$= \tfrac{1}{2}\log 2 - \frac{5\pi}{4} - 2h\pi + i\left(\tfrac{1}{2}\log 2 + \frac{5\pi}{4} + 2h\pi\right).$$

To bring $(-1-i)^{1+i}$ in the standard form of a complex number we make use of formulas (6.02) and (6.011). Thus we find that $(-1-i)^{1+i}$ has the denumerable infinitude of values represented by the expression

$$e^{\frac{1}{2}\log 2-\frac{5\pi}{4}-2h\pi}\left[\cos\left(\tfrac{1}{2}\log 2+\frac{5\pi}{4}\right)+i\sin\left(\tfrac{1}{2}\log 2+\frac{5\pi}{4}\right)\right],$$

in which h is an arbitrary integer.

We are now prepared to deal with the problem of this chapter in its most general form: If α and β are given complex numbers, $\alpha \neq 1$, to determine a complex number c, such that $\alpha^c = \beta$. For this equation is equivalent to the equation $c \log \alpha = \log \beta$, so that we find $c = \dfrac{\log \beta}{\log \alpha}\cdot$ By means of the interpretation which Definition XXVII gives to these symbols, we can write our result in the form

$$c = \frac{\log m(\beta) + i\,\theta(\beta)}{\log m(\alpha) + i\,\theta(\alpha)},$$

in which $\theta(\beta)$ and $\theta(\alpha)$ each have a denumerable infinitude of values. The quotient of two complex numbers which appears in this formula can be reduced to a complex number; for we know that the set C_2 of complex numbers is closed with respect to all the rational operations, in particular with respect to division (compare 45, pp. 80–83). We are therefore justified in the statement of the following theorem.

Theorem XVIII. The field of complex numbers is closed with respect to the second inverse of involution, i.e. with respect to the operation of taking a logarithm.

The result which we have obtained above can be given a slightly different emphasis. In accordance with Definition XXVI (p. 115), we can say that if $\alpha^c = \beta$, then c is the logarithm of β with respect to the base α. Thus we obtain the statement contained in

Theorem XIX. If α and β are arbitrary complex numbers, $\alpha \neq 1$, then the logarithm of β with respect to the base α is equal to the quotient of any value of the *natural* logarithm of β by any value of the natural logarithm of α.

It is frequently convenient to single out one value of the amplitude of a complex number and one value of its logarithm. We lay down the following definition:

Definition XXIX. The *principal value of the amplitude* of a complex number α is the value of $\theta(\alpha)$ which lies between 0° and 360° (including 0°, excluding 360°), i.e., in radian measure, the value of $\theta(\alpha)$ which lies between 0 and 2π (including 0, but not 2π). The *principal value of the natural logarithm* of α is equal to $\log m(\alpha) + i \times$ the principal value of $\theta(\alpha)$.

The principal value of $\log \alpha$ is designated by $\operatorname{Log} \alpha$, so that $\log \alpha = \operatorname{Log} \alpha + 2n\pi i$.

Let us now look at a few examples.

Example 1. To solve the equation $(-1)^c = -2$, we have to determine $\log(-1)$ and $\log(-2)$. We know that $m(-1) = 1$, $\theta(-1) = \pi + 2n\pi$; $m(-2) = 2$, $\theta(-2) = \pi + 2h\pi$. Hence $\operatorname{Log}(-1) = \pi i$ and $\operatorname{Log}(-2) = \log 2 + \pi i$. Consequently

$$c = \frac{\log(-2)}{\log(-1)} = \frac{\log 2 + (2h+1)\pi i}{(2n+1)\pi i} = \frac{2h+1}{2n+1} - \frac{i \log 2}{(2n+1)\pi}$$

n and h being arbitrary integers. These values of c are also the values of $\log_{-1}(-2)$. For $h = n = 0$, we find $1 - \dfrac{i \log 2}{\pi}$.

To verify the last result, we notice that

$$(-1)^{1-\frac{i\log 2}{\pi}} = (-1) \cdot e^{-\frac{i\log 2}{\pi} \cdot \operatorname{Log}(-1)} = -e^{-\frac{i\log 2}{\pi} \cdot i\pi} = -e^{-i^2 \log 2} = -e^{\log 2} = -2.$$

Example 2. Without difficulty we find that: $\operatorname{Log} i = i \times$ principal value of $\theta(i) = \dfrac{i\pi}{2}$, and hence that $i = e^{\frac{i\pi}{2}}$. This result is immediately verified by means of Euler's formula (6.01).

66. A few side views. Before we proceed to a rest period we must look at a few sights which we passed by on our final dash to the goal.

The field of complex numbers has a part which is isomorphic with the set of positive real numbers (see *46*) with respect to addition, division and multiplication, for it contains the complex numbers $(a, 0)$ in which a is a positive real number; the amplitude of these numbers is 0. But in the light of the discussion of *51* (see p. 92), the amplitude of these complex numbers has not only the value 0, but also 360°, $2 \cdot 360°$, . . . $n \cdot 360°$, . . . (or 2π, 4π, . . . $2n\pi$, . . .). In the terminology of Definition XXIX, the principal value of their amplitude is 0; moreover if α is a complex

number of this type, i.e. if $\alpha = (a, 0)$ where a is a positive real number, $m(\alpha) = a$. Consequently $\text{Log}\ \alpha = \log a$, and $\log \alpha = \log a + 2n\pi i$. That is to say, the isomorphism between the system consisting of the complex numbers of the form $(a, 0)$ with the rational operations except subtraction, and the system {positive real numbers, rational operations except subtraction} (see *46*) is preserved under the operation of taking the logarithm, provided that for the complex numbers $(a, 0)$ only the principal value of the logarithm is considered. In other words the system {complex numbers of form $(a, 0)$ with a positive, rational operations except subtraction, principal value of logarithm} is isomorphic with the system {positive real numbers, rational operations except subtraction, logarithm}.

The process which led us to Theorem XIX and enabled us to find $\log_\alpha \beta$ in terms of the natural logarithms of α and β, simplifies a good bit when we restrict ourselves to positive real numbers; it leads to a result of importance in the calculation of logarithms of such numbers. Suppose that a and b are positive real numbers and $\alpha = (a, 0)$, $\beta = (b, 0)$. Then, as we have just seen, $\text{Log}\ \alpha = \log a$ and $\text{Log}\ \beta = \log b$. Consequently, the real number c, which solves the equation $a^c = b$, is equal to $\dfrac{\log b}{\log a}\cdot$ In contrast with the result stated in Theorem XIX, we have now a single real number for $\log_a b$.

Theorem XX. If a and b are positive real numbers, then the logarithm of b with respect to the base a is equal to the quotient of the natural logarithm of b divided by the natural logarithm of a.

Now it so happens, for reasons which are some distance removed from our path (they lie rather near the source of Euler's formula (6.01) and would be uncovered by an expedition into that region), that the *natural* logarithms of positive real numbers can be more readily calculated than logarithms with respect to any base other than e. On the other hand for many applications, the *common* logarithms (i.e. logarithms with respect to the base 10, see p. 119) are the most useful. Theorem XX now shows the way in which we can pass from the natural logarithms to the common logarithms; for, if we put $a = 10$, we obtain the following corollary.

Corollary. If b is a positive real number, the common logarithm of b is equal to the quotient of its natural logarithm by the natural logarithm of 10:

$$\log_{10} b = \frac{\log b}{\log 10}.$$

Consequently a table of common logarithms is obtainable from a table of natural logarithms by dividing all the entries in the latter by the natural logarithm of 10. The only remaining mystery is how to construct a table of natural logarithms — it is one of the problems for the technician. We must proceed with some exercises.

67. A last look around.

1. Change the following angle measurements to radians: $225°$, $330°$, $48°$, $150°$, $216°$, $420°$.

2. Change the following radian measurements of angles to degrees: $\frac{2\pi}{3}, \frac{3\pi}{5}, \frac{11\pi}{6}, \frac{7\pi}{4}, \frac{3\pi}{2}, \frac{15\pi}{8}, \frac{5\pi}{6}, \frac{2\pi}{9}$.

3. Determine the natural logarithms of the following complex numbers: $-5, i, -i, 1+i, -\sqrt{3}+i, 1-i, -1-i\sqrt{3}, 4, -\sqrt{3}-i$.

4. Interpret the symbol i^i.

5. Show that $i^{-i} = (-i)^i$, in a sense similar to the one in which Theorem XVII, (1) was proved for complex numbers.

6. Determine Log $(1+i)$, Log $(\sqrt{3}-i)$.

7. Determine $\log_{-1} i$.

8. Prove that if α and β are complex numbers whose modulus is unity, then all values of $\log_\alpha \beta$ and of $\log_\beta \alpha$ are real numbers.

9. Determine $\log_i (1+i)$.

10. Prove Theorem XX directly from Theorem XVII, without use of Theorem XIX.

11. Prove that if a, b and c are positive real numbers, then $\log_b a = \log_c a \cdot \log_b c$.

12. Deduce from 11 the fact that $\log_b a \cdot \log_a b = 1$.

13. Prove that Theorem XVII, (2) is valid for complex numbers.

14. Derive de Moivre's formula (see *56*, 13, p. 101) from Euler's formula (6.01). Does this derivation constitute a proof?

15. Is it true that Log pq is equal to Log p + Log q, when p and q are arbitrary complex numbers?

16. Is it true that Log p^k is equal to k Log p, when p is an arbitrary complex number and k is a positive integer?

17. Discuss the question of 16 for the case that k is a proper fraction.

18. Solve the equation $x^4 = \cos 20° + i \sin 20°$.

19. Solve the equation $x^6 = i$.

20. Discuss the expression $q^{\log_q p}$ if p and q are arbitrary complex numbers.

CHAPTER VII

AN UNCONVENTIONAL EPISODE

The human mind is not exhausted; it searches and continues to find so that it may know that it can find indefinitely, and that only laziness can limit man's knowledge and his inventions. — *Bossuet.*

68. Looking at the map. We return once more to the scene of our first journey in order to pick up a trail that will lead us to new experiences. It branches off slightly beyond the point from which the second excursion started, viz. at the place where we observed that, since $q = f\sqrt{r_1^2 - p_1^2}$, q must have the factor f. We had glimpses of our new trail at some other points. For instance, in the proof of Lemma 3 (see p. 9), we had found that $2t = k(u + v)$ and concluded from this that all the factors of k, with the possible exception of a factor 2, must be factors of t. The argument used on pages 51 and 55 to establish the non-existence of rational solutions of the equations $x^2 = 2$ and $x^3 = 9$ respectively is of the same type. Carried over into the domain of natural numbers, we would, since 2×1134 is equal to 42×54, infer that all the factors of 42, a factor 2 perhaps excepted, must also be factors of 1134. In this particular instance the assertion can readily be tested; any one who makes the test finds that 42 is indeed a factor of 1134. But what is the fundamental reason which underlies this fact? Neither 1134, nor 42, nor 54 are prime numbers. Indeed we find that $1134 = 2 \times 3 \times 3 \times 3 \times 3 \times 7$, so that $2 \times 1134 = 2^2 3^4 7$; and, $42 = 2 \times 3 \times 7$ and $54 = 2 \times 3 \times 3 \times 3$, so that 42×54 is also equal to $2^2 3^4 7$.[1] Since now 2×1134 and 42×54 give rise to exactly the *same products of prime factors*, all the prime factors of 42 must occur among the prime factors of 2×1134 and therefore, barring perhaps a single factor 2, among those of 1134. If it had been possible to factor 2268 into prime factors in more than one way, our conclusion could not have been drawn. But this, every one will say, is impossible. Why is this impossible? Don't we have

[1] Observe that the commutative and associative laws of multiplication play important parts in the derivation of these statements (compare p. 42).

two products, 2×1134 and 42×54, which are both equal to 2268? and are not the products 4×567, and 36×63, and 21×108 also equal to this same number 2268? Yes, that is indeed true; but these products are not products of prime numbers. The contention is therefore that 2268 *can be factored into prime factors in only one way*. But on what is this contention based?

At this point the dialogue between the persistent questioner and the stubborn defender has reached the stage at which an appeal to authority is probably imminent. The answer to the last question is very likely to be: the books say so; and if you want to know why they say so, look up how they justify themselves. It is indeed true that a good many books on algebra make the statement, with more or less satisfactory justification, that every natural number can be factored into prime factors in only one way. Here it is to be understood that the factor set (compare pp. 33 and 35) is the set of natural numbers. The statement is usually called the *unique factorization theorem* for natural numbers. There is not much use in repeating any of the usual arguments brought forward to support the contention. Instead we shall try to put the reader in a state of mind in which he may feel the need for a proof of this unique factorization theorem in the set of natural numbers. To this end we shall journey into a realm, in which there is *no* unique factorization theorem. After the completion of this excursion we may think it desirable to make sure that the set of natural numbers with which so much of our experience is intimately concerned does not lack so essential a guaranty of security and safety. Not until after we have experienced the difficulties which arise as a result of its absence, can we properly appreciate the value and significance of this guaranty. It is a hackneyed, but none the less sound piece of warning to the young, that the good things of life are too rarely appreciated while they are within reach; we do not know what it is to have a sound body until we are tortured by aches and pains. And now we are off for the wilderness once more.

69. A new setting for familiar things. We have become thoroughly acquainted by now with the rational numbers and with some of their important properties (the forgetful reader can refresh his memory on pages 4, 27, 38–40 and 47–49). The fact that every rational number is of the form $\frac{p}{q}$, in which p and q are relatively prime integers, positive, zero or negative, except that

$q \neq 0$, can be expressed as follows: Every rational number is a solution of an equation of the form

$$(7.1) \qquad qx - p = 0,$$

in which p and q are relatively prime integers, $q \neq 0$.

Moreover every equation of this form has exactly one solution, viz. the rational number $\frac{p}{q}$. In the language with which we have become familiar in Chapter II, this statement becomes: The set of rational numbers is equivalent to the set of equations of type (7.1). We have furthermore the following special case: The set of rational integers is equivalent to the set of equations $x - p = 0$.

From this simple beginning we are led to a natural generalization. The set of equations of type (7.1), in which the variable x occurs to the first power at most, is called the set of *algebraic equations of the first degree*, or *the set of linear algebraic equations*. It is contained as a proper part in the set of equations

$$(7.2) \qquad a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n = 0,$$

in which n is a natural number, and $a_0, a_1, \ldots a_{n-1}, a_n$ is a set of relatively prime integers, positive, zero or negative, except that $a_0 \neq 0$. Equation (7.2) is called the *general algebraic equation*.

The use of letters with subscripts, like a_0, a_1, a_2 etc., is already familiar from earlier work. It is particularly convenient for the general algebraic equation, in which the degree n is not definitely specified; in every term the sum of the subscript of the coefficient and the exponent of the variable is n. Thus, for example, in an equation of degree 20 (the special case in which $n = 20$), the term involving x^9 would have the form $a_{11}x^9$. In the general case the terms can not all be written down; the dots serve to indicate the missing terms, from a_2x^{n-2} down to $a_{n-1}x$, of which the total number is left unspecified. It is of course immaterial whether we use x, or y, or any other letter for the variable, and whether the coefficients are represented as in (7.2) or by some other letter. We shall understand when we speak of "an equation of type (7.2)" that we mean an equation of the form indicated above, in which the coefficients are arbitrary relatively prime integers, with the restriction $a_0 \neq 0$.

By analogy with the relation established above between the rational numbers and equations of type (7.1), we lay down now the following generalization of rational numbers and rational integers.

Definition XXX. An *algebraic number* is any number which satisfies an equation of type (7.2).

Definition XXXI. An *algebraic integer* is any number which satisfies an equation of type (7.2) in which $a_0 = 1$.

We observe in the first place that the set of rational numbers is contained in the set of algebraic numbers, and that the set of rational integers is a subset of the set of algebraic integers. The term "rational integers" is used to designate the integers in the field R of rational numbers to distinguish them from the algebraic integers with which they frequently enter in combination. But what else is covered by Definitions XXX and XXXI?

70. Algebraic and transcendental numbers — squaring the circle. Let us consider the special case of equation (7.2) in which $n = 2$, i.e. the equation

$$(7.3) \qquad a_0x^2 + a_1x + a_2 = 0,$$

in which a_0, a_1, a_2 are three relatively prime integers and $a_0 \neq 0$. This equation is called the *general quadratic equation*, or equation of *the second degree;* it is discussed in every book on elementary algebra. In every such book (and they are so numerous that no special reference is necessary), the reader will find a statement and justification of the fact that there are two numbers which satisfy this equation, viz. the numbers r_1 and r_2, called the roots of the quadratic equation, which can be calculated from the following formulas:

$$(7.4) \quad r_1 = \frac{-a_1 + \sqrt{a_1^2 - 4a_0a_2}}{2a_0} \text{ and } r_2 = \frac{-a_1 - \sqrt{a_1^2 - 4a_0a_2}}{2a_0}.$$

What sort of numbers are these? In the first place, whatever else they may be, in accordance with Definition XXX, they are algebraic numbers. Do they belong to any of the categories of numbers which we have met on previous parts of our journey? Clearly both r_1 and r_2 contain the term $-\frac{a_1}{2a_0}$; this, being the quotient of two integers, is surely a rational number. As to the second term, it contains the expression $\sqrt{a_1^2 - 4a_0a_2}$. Now $a_1^2 - 4a_0a_2$ is surely an integer; it may be positive, zero or negative. In any case, in virtue of Theorem XII (see p. 91), there always exists a complex number X, such that $X^2 = a_1^2 - 4a_0a_2$; consequently $\sqrt{a_1^2 - 4a_0a_2}$ is a complex number which may or may

not be real. Since the rational numbers are contained among the complex numbers and since the set of complex numbers is closed under division and addition (i.e. since the quotient and the sum of two complex numbers are again complex numbers), it should be evident that r_1 and r_2 are complex numbers.

This conclusion shows us once more that the labor we went through to gain the greater freedom which was acquired in Chapter V was indeed worth while; the extension of our field of operations which it made possible enables us to solve the equation (7.3) i.e. the general quadratic equation without leaving the field of complex numbers. We see moreover that the set of algebraic numbers contains, besides the rational numbers, at least certain other kinds of complex numbers, so that the generalization from the set of rational numbers to the set of algebraic numbers as defined in Definition XXX represents indeed an extension.

But we can make an even more significant statement. Not only the general quadratic equation (7.3), but *every* algebraic equation of type (7.2) can be solved within the field of complex numbers; in other words all algebraic numbers are complex numbers, or again, the set of algebraic numbers is contained in the set of complex numbers. We have to be satisfied here with a mere statement of this fact; a proof is found in books dealing with fundamental questions of algebra and in books on the *Theory of Functions of a Complex Variable*. It can be established without difficulty after we have established a theorem, known as the *fundamental theorem of algebra*, which says that every equation of type (7.2) can be satisfied by at least one complex number. To determine such a complex number for an equation of higher degree than the second (i.e. for an equation (7.2) in which $n > 2$) is not always a simple matter; it is a technical question which is dealt with in the further study of algebra. We return now to equation (7.3) and its roots (7.4).

If $a_1^2 - 4a_0a_2$ should happen to be the square of an integer, r_1 and r_2 as defined by (7.4) are seen to be rational numbers. If it is positive, but not the square of an integer, they are real numbers (compare Theorem IX, p. 69, and p. 84). If it is negative, then $\sqrt{a_1^2 - 4a_0a_2}$ is a normal number, and r_1 and r_2 are non-real complex numbers. Thus we see that the set of algebraic numbers contains numbers of each of the types discussed in Chapters III, IV and V.

The question naturally arises whether all real numbers are also algebraic numbers. The answer to this question is in the negative; as we know (compare pp. 51, 102) a single example suffices to substantiate the negative answer. The number π is such an example; it is a real number but not algebraic. This fact is here stated without proof; unfortunately, the proof is not accessible to us. It is of comparatively recent date and utilizes a number of ideas which are beyond our range. Once more we have reached a point in our travels from which we can see an alluring mountain top that we can not undertake to climb. Let us walk around its base for a while.

One of the problems with which Greek mathematicians occupied themselves a great deal from the fifth century B.C. onwards is that of "squaring the circle," i.e. the construction of a square equal in area to the area of a given circle.[1] The solution of this problem requires the construction of a line equal to $a\sqrt{\pi}$, in which a is given. The essential step in this construction is the construction of a line equal to π; for a line of length $a\sqrt{\pi}$ can easily be drawn once the construction of a line equal to π has been accomplished.[2] Now it should be clear that a problem of this sort, in which we are asked to construct or to do something, lacks definiteness until it is specified what tools are to be at our disposal. To make a radio set is one problem when the only things we are allowed to use are hammers and nails, but quite a different one if we can command the resources of a modern electrical laboratory. The task may be impossible in the one case, and entirely within our possibilities in the other. To play a Chopin étude on the piano when both hands can be used is a different undertaking from performing the same étude with the left hand alone. In the same way the problem of "squaring the circle" becomes a definite one only when the means that can be used for it have been specified. It is surprising that it took so many years to clarify this point and to bring out the fact that the interesting and significant formulation of the problem is "*to construct a line of length π by means of a straight edge*[3] *and a pair of compasses only.*" This problem turned out to be a hard one to solve.

[1] See e.g. Heath, *A Manual of Greek Mathematics*, pp. 139–147.

[2] Compare pp. 161–164; see also pp. 73–75.

[3] We are to understand here and in the sequel that the straight edge is not marked off.

It is a rather simple matter to show that whenever a line can be constructed from a given unit line by the use of a straight edge and a pair of compasses only, then the length of that line is expressed by a number which is algebraic. Not until 1882 did the German mathematician Ferdinand Lindemann (1852–) succeed in proving that the squaring of the circle by means of straight edge and compasses could not be accomplished, by showing that the number π *is not an algebraic number*, i.e. that there is no equation of type (7.2) which has the number π as one of its roots. This was a prodigious accomplishment which brought to a close the study of one of the famous problems of antiquity. The story of this problem forms a very remarkable chapter in the history of mathematics — a few references must suffice at this point.[1]

It is perhaps not surprising that many people who lacked the knowledge necessary to appreciate the difficulty of the problem, attempted its solution before 1882. That such useless waste of energy continued even after that time up to the present is an illustration of the slowness with which knowledge, and especially knowledge of mathematics, penetrates to the mass of mankind. "Circle-squarers" appear to be an over-fertile section of the human family; neither the advance of knowledge nor the sting of persistent ridicule have succeeded in wearing down their ambition.[2]

The work of Lindemann followed the methods used by the French mathematician Charles Hermite (1822–1901) who proved in 1873 that the number e (see p. 120) *is not an algebraic number.* In all this work an important rôle is played by a general theorem, proved in some of the books to which reference has been made, viz. the theorem that if $e^x = y$, then x and y can not both be algebraic numbers, except when $x = 0$ and $y = 1$. From this theorem and the relation $e^{\pi i} = -1$, mentioned on page 122, it follows that πi is not an algebraic number, and hence that π is not an algebraic number.[3]

These results show that there are at least two real numbers,

[1] See E. W. Hobson, *Squaring the Circle;* J. W. Young, *Fundamental Concepts of Algebra and Geometry*, pp. 120 and 191; F. Klein, *Famous Problems of Elementary Geometry*, translated by Beman and Smith; D. E. Smith, *Source Book of Mathematics*, p. 99; *Encyclopædia Britannica*, 11th edition, vol. 6, pp. 384–387; Hessenberg, *Transzendenz von e und π;* J. W. A. Young, *Monographs on Topics of Modern Mathematics,* Chapter IX.

[2] Compare the delightful book *A Budget of Paradoxes*, by Augustus de Morgan (1806–1871), the father of the well-known novelist Joseph de Morgan, author of *Joseph Vance*, *Alice-for-short*, and many other fascinating books.

[3] Compare the first section of the book by Hessenberg, quoted above.

viz. π and e, which are not algebraic. We shall see presently that of the non-denumerable set of real numbers which forms the continuum (see p. 30) only a denumerable subset is algebraic, so that, in a sense which we should now readily understand, there are more real numbers which are not algebraic than there are algebraic real numbers. Non-algebraic real numbers are designated by a special name.

Definition XXXII. A real number which is not an algebraic number is called a *transcendental real* number. A complex number which is not algebraic is called a *transcendental* number.

The numbers e and π are therefore transcendental real numbers. The theorem mentioned above can now be stated as follows: If $e^x = y$, and one of the two numbers x and y is algebraic, then the other one is transcendental, except when $x = 0$ and $y = 1$.

We have now seen that some real numbers are algebraic numbers; also, that there are algebraic numbers which are not real, but that all algebraic numbers are contained in the set of complex numbers. The algebraic numbers constitute a set of complex numbers, which overlaps the different divisions of this field with which we have already become acquainted. In studying them we roam over a realm of the field of complex numbers whose boundaries cut across the boundaries which our former study has erected (and torn down!).

71. Fields of algebraic numbers. We turn now to the consideration of some special classes of algebraic numbers.

(a) All complex numbers of the form $a + bi$, in which a and b are rational numbers, are certainly algebraic. For if we suppose $a = \dfrac{p}{q}$ and $b = \dfrac{p_1}{q_1}$, where p, q, and p_1, q_1 are pairs of relatively prime integers, and if we put $x = \dfrac{p}{q} + \dfrac{p_1 i}{q_1}$, we find that $qq_1x = pq_1 + p_1qi$. Consequently $(qq_1x - pq_1)^2 = p_1^2q^2i^2 = -p_1^2q^2$, so that $a + bi$ satisfies the equation

$$(7.5) \qquad q^2q_1^2x^2 - 2pqq_1^2x + p^2q_1^2 + p_1^2q^2 = 0;$$

but this is an equation of type (7.3). Therefore, in accordance with Definition XXX, $a + bi$ is indeed an algebraic number.

We know moreover, if we recall Definition IX (see p. 45) and 27, 11 that the set of numbers $a + bi$, in which a and b are rational numbers, constitutes a field. Those numbers of the field

which are algebraic integers, are briefly called the integers of the field, a name which we will justify presently. Which are the integers of the field of numbers $a + bi$?

It follows from Definition XXXI and equation (7.5) that $\frac{p}{q} + \frac{p_1 i}{q_1}$ is an algebraic integer if and only if the integers $2pqq_1^2$ and $p^2q_1^2 + p_1^2q^2$ are both divisible by the integer $q^2q_1^2$, the factor set consisting of the integers.[1] This requires that $\frac{2p}{q}$ and $\frac{p^2}{q^2} + \frac{p_1^2}{q_1^2}$ be integers. From the first of these conditions follows, since p and q have no factor in common, that q must be either 2 or 1. In the former case $q^2 = 4$, so that, if $\frac{p^2}{q^2} + \frac{p_1^2}{q_1^2}$ is to be an integer, q_1^2 must also equal 4, and $p^2 + p_1^2$ must be divisible by 4. Moreover p and p_1 must be odd; hence, representing p and p_1 in the form $2n + 1$ and $2n_1 + 1$ respectively, $(2n + 1)^2 + (2n_1 + 1)^2$ must be divisible by 4. But $(2n + 1)^2 + (2n_1 + 1)^2 = 4(n^2 + n_1^2 + n + n_1) + 2$ which obviously leaves a remainder 2 if divided by 4. Therefore the whole stack of blocks built up on the assumption that $q = 2$ tumbles down; and we have to conclude that $q = 1$.

But, if $q = 1$ and $\frac{p^2}{q^2} + \frac{p_1^2}{q_1^2}$ is an integer, say k, then q_1 must also be equal to 1, for otherwise we would have to conclude that the irreducible fraction $\frac{p_1^2}{q_1^2}$ is equal to the integer $k - p^2$. A glance at equation (7.5) shows that if, conversely, $q = q_1 = 1$, then $\frac{p}{q} + \frac{p_1 i}{q_1}$ is certainly an algebraic integer. The upshot of our discussion is therefore that $\frac{p}{q} + \frac{p_1 i}{q_1}$ is an algebraic integer if and only if $q = q_1 = 1$; i.e. the integers of the field of numbers $a + bi$, in which a and b are

[1] The argument is not complete at this point, since, if $2pqq_1^2$ and $p^2q_1^2 + p_1^2q^2$ are not both divisible by $q^2q_1^2$, we can only affirm that the *quadratic* equation which $\frac{p}{q} + \frac{p_1 i}{q_1}$ satisfies is not of the form required for an algebraic integer, but not that this number can not satisfy an equation of this character of higher degree. A similar remark has to be made in parts (b) and (c) below. It would take us too far afield to complete the argument. The reader is referred to books on algebraic numbers, see, e.g., L. W. Reid, *The Elements of the Theory of Algebraic Numbers*, or E. Landau, *Einführung in die elementare und analytische Theorie der algebraischen Zahlen und der Ideale.*

rational numbers, is the set of numbers $p + p_1 i$ in which p and p_1 are integers.

The name "integers of the field" which we have attached to the "algebraic integers" of the field of the complex numbers $a + bi$, in which a and b are rational numbers, now finds its justification. We have just seen that if and only if a and b are rational integers, $a + bi$ will be an integer of the field. As "integers" they should certainly constitute a set that is closed with respect to addition and multiplication. This is indeed the case, for we know from Chapter V (see pp. 79, 82, 86) that $(a + bi) + (a_1 + b_1 i) = (a + a_1) + (b + b_1)i$, and that $(a + bi)(a_1 + b_1 i) = (aa_1 - bb_1) + (ab_1 + ba_1)i$; from the closure properties of the set of rational integers, the desired closure of the integers of the field follows therefore immediately.

(b) If a and b are rational numbers, all numbers of the form $a + b\sqrt{5}$ are algebraic. For, if we follow the procedure just used in (a), we find that such a number satisfies the equation

$$q^2 q_1^2 x^2 - 2pqq_1^2 x + p^2 q_1^2 - 5p_1^2 q^2 = 0. \tag{7.6}$$

Furthermore these numbers form a field (compare 27, 10). Postponing the justification of the name till later, we designate the algebraic integers which satisfy equation (7.6) as the "integers of the field $a + b\sqrt{5}$." The search for the integers requires a more elaborate expedition in this than in the previous case.

We see that if $\dfrac{p}{q} + \dfrac{p_1\sqrt{5}}{q_1}$ is to be such an integer, then both $-2pqq_1^2$ and $p^2 q_1^2 - 5p_1^2 q^2$ must be divisible by $q^2 q_1^2$. This will surely be the case if q and q_1 are both equal to 1. But it also happens, for instance, if $p = p_1 = 1$ and $q = q_1 = 2$; equation (7.6) then becomes $16x^2 - 16x - 16 = 0$, which reduces to $x^2 - x - 1 = 0$. It is not difficult to see that whenever $q = q_1 = 2$ and $p = 2n + 1$, $p_1 = 2n_1 + 1$ (i.e. p and p_1 odd), $\dfrac{p}{q} + \dfrac{p_1\sqrt{5}}{q_1}$ will be an algebraic integer. For in this case $q^2 q_1^2 = 16$, $2pqq_1^2 = 16(2n + 1)$, and $p^2 q_1^2 - 5p_1^2 q^2 = 4(4n^2 + 4n + 1) - 5 \cdot 4 \cdot (4n_1^2 + 4n_1 + 1) = 16(n^2 + n - 5n_1^2 - 5n_1) + 4 - 20 = 16(n^2 + n - 5n_1^2 - 5n_1 - 1)$, so that equation (7.6) reduces to the form $x^2 - (2n + 1)x + n^2 + n - 5n_1^2 - 5n_1 - 1 = 0$. Therefore, in accordance with Definition XXXI, the field whose elements are $a + b\sqrt{5}$

(a and b rational numbers) contains as integers not merely the numbers $p + p_1\sqrt{5}$, in which p and p_1 are rational integers, but also numbers of the form $\frac{p}{2} + \frac{p_1\sqrt{5}}{2}$ where p and p_1 are odd rational integers.

And now we will show that this field contains no other integers. For if $-2pqq_1^2$ is to be divisible by $q^2q_1^2$, then $\frac{2p}{q}$ must be an integer; and hence, since p and q are relatively prime, if q is not equal to 1, it must equal 2. But then the condition that $p^2q_1^2 - 5p_1^2q^2$ be divisible by $q^2q_1^2$ becomes the condition that $p^2q_1^2 - 20p_1^2$ be divisible by $4q_1^2$, i.e. that $p^2q_1^2 - 20p_1^2 = 4q_1^2k$, where k is an integer. If we put this equation in the form $(p^2 - 4k)q_1^2 = 20p_1^2$, it follows from the unique factorization theorem [1] that, since $20p_1^2$ has the obvious factor 4, its left hand side must be divisible by 4. Now, since p and q are relatively prime and $q = 2$, p is odd; hence p^2 and also $p^2 - 4k$ is odd, so that $p^2 - 4k$ has no factor in common with 4. Therefore q_1^2 must contain the factor 4 and q_1 the factor 2. Furthermore q_1 can not have any other factor. For if it did q_1^2 would contain the same factor twice; but this factor could not occur in p_1^2, since p_1 and q_1 are relatively prime, and 20 has no other square factor besides 4. If, on the other hand, $q = 1$, then it would follow from an argument, similar to the one we have just used, that q_1 must also equal 1. Therefore if q and q_1 are not both equal to 1, they must both be equal to 2, if $\frac{p}{q} + \frac{p_1\sqrt{5}}{q_1}$ is to be an integer of the field. The integers of this field can then be represented in the form $\frac{p}{2} + \frac{p_1\sqrt{5}}{2}$, where p and p_1 are either both even, or both odd, i.e. of equal parity (see p. 10; compare also 72, 6).

(c) Every number of the form $a + b\sqrt{-6}$ satisfies an equation of the form

$$(7.7) \qquad q^2q_1^2x^2 - 2pqq_1^2x + p^2q_1^2 + 6p_1^2q^2 = 0,$$

where p, q and p_1, q_1 are pairs of relatively prime integers and $a = \frac{p}{q}$,

[1] This is not the only point in the present section at which this theorem is used; the reader will do well to point out the other places in the argument at which it plays a part and at which it has not been mentioned.

$b = \frac{p_1}{q_1}$. These numbers form a field. The integers of the field are those numbers $\frac{p}{q} + \frac{p_1\sqrt{-6}}{q_1}$ for which $\frac{2p}{q}$ and $\frac{p^2}{q^2} + \frac{6p_1^2}{q_1^2}$ are integers; as in (a) we find that q must be either 2 or 1.

Suppose $q = 2$; then $\frac{p^2}{4} + \frac{6p_1^2}{q_1^2}$ must be equal to an integer, say k, i.e. $p^2q_1^2 + 24p_1^2 = 4kq_1^2$, or

$$q_1^2(4k - p^2) = 24p_1^2. \tag{7.8}$$

Now we proceed as in (b); since $24p_1^2$ contains the factor 4, the left hand side of (7.8) must also contain the factor 4. But if $4k - p^2$ were even, p^2 would have to be even, hence p even so that p and q would not be relatively prime. The unique factorization theorem leads then to the conclusion that q_1^2 must contain the factor 4 and q_1 the factor 2. It might be thought that q_1 can contain other factors besides 2; if it did these would have to be factors either of p_1^2 or of 6, i.e. 2 or 3. But q_1 has no factors in common with p_1; and q_1^2 could not have either of the factors 2 and 3 without having two such, one of which would then have to appear also in p_1. Therefore if $q = 2$, then $q_1 = 2$. Moreover p and p_1 have to be odd *and* have to satisfy equation (7.8), which now takes the form $4k = p^2 + 6p_1^2$. These conditions are evidently incompatible since, if p were odd, p^2 would also be odd, hence $p^2 + 6p_1^2$ odd and not equal to $4k$, which is even.

Summa summarum, q can not be equal to 2; there is no escape from the conclusion $q = 1$. The reader should now have no difficulty in showing that q_1 must then also be 1. After he has shown this, he can conclude that the only integers of the field $a + b\sqrt{-6}$ (a and b rational) are the numbers $p + p_1\sqrt{-6}$, in which p and p_1 are rational integers.

These examples of fields of algebraic numbers are instances of what are called *quadratic* fields — their elements all satisfy quadratic equations. In the theory of algebraic numbers, which has been developed very extensively in recent times, one considers *cubic* fields, *quartic* fields, etc. It should be evident that there is enormous scope for discovery in these domains. Many questions suggest themselves, of which only a few have been hinted at. What are the *unit* and *zero* elements in these fields of algebraic numbers?

Evidently the rational numbers 0 and 1 belong to every such field. It is seen at once that they have in these fields the properties which they are known to have in the field R of rational numbers. But, are these the properties by which they should be characterized in algebraic fields? One of the properties of the numbers $+1$ and -1 in the field of rational numbers is the fact that they are the only integers whose reciprocal is also an integer. It is this property which is used to define the *units* in a quadratic field. It is clear that $+1$ and -1 are then units of every such field. Moreover, it is readily seen that such a field may have other units. We have but to observe that the integers $+i$ and $-i$ of the field $a + bi$ are reciprocals of each other; that the product of the integers $\sqrt{5} + 2$ and $\sqrt{5} - 2$ of the field $a + b\sqrt{5}$, and the integers $2 + \sqrt{3}$ and $2 - \sqrt{3}$ of the field $a + b\sqrt{3}$ have a product equal to $+1$. A fuller discussion of units in quadratic fields will be found in the books mentioned in the footnote on page 138.

In our search for the integers of quadratic fields we have not obtained any general results so far. While in the field $a + b\sqrt{5}$, all numbers of the form $\frac{p}{2} + \frac{p_1\sqrt{5}}{2}$ are integers, whenever p and p_1 are rational integers of the same parity, the field $a + b\sqrt{-6}$ has as integers only the numbers $p + p_1\sqrt{-6}$, p and p_1 being arbitrary rational integers. The reader who has followed the argument of this section should be able to account for this difference; he will have a chance at the next resting place to try his strength at the problem (compare *72*, 15 and 16). If he finds himself unequal to the task, he had better go back over the path we have followed to search for a clue.

After a little refreshment and recreation we shall go a little further in the study of quadratic fields; we must push ahead far enough to discover the horrifying fact that in these fields the unique factorization theorem does not always hold.

72. Going afield.

1. Write in full the general algebraic equations of the 3rd, 4th, 5th and 6th degrees.

2. Show that $\sqrt{5 - 2\sqrt{3}}$ is an algebraic number and set up an algebraic equation which this number satisfies.

3. Prove that the set of algebraic numbers which are roots of quadratic equations is a denumerable set.

4. Show for each of the following numbers that they are algebraic, and exhibit for each of them an algebraic equation which it satisfies:

(a) $2 + \sqrt[3]{-2}$, (b) $\sqrt{-3 + \sqrt[3]{5}}$, (c) $\sqrt{2} - \sqrt[3]{3}$.

5. Determine the integers in the field of numbers of the form $a + b\sqrt{3}$, where a and b are rational numbers.

6. Show that the integers of the field $a + b\sqrt{5}$ form a set that is closed with respect to addition and multiplication (compare *71*, (a), (b)).

7. Solve the corresponding problem for the field $a + b\sqrt{-6}$ (compare *71*, (c)).

8. Show that the set of equations (7.2) is not effectively enlarged if the coefficients $a_0, a_1, \ldots a_n$ are chosen from the set of rational numbers.

9. Prove that the natural logarithms of the positive integers except 1 are all transcendental numbers.

10. Prove that all powers of e with non-zero integral exponents are transcendental numbers.

11. Determine the integers of the field whose elements are the numbers of the form $a + b\sqrt{-5}$, a and b rational.

12. Justify the use of the word integer for the numbers determined in the preceding problem.

13. Prove that in the field whose elements are the numbers of the form $a + b\sqrt{13}$, a and b rational, all numbers of the form $\dfrac{p}{2} + \dfrac{p_1\sqrt{13}}{2}$ are integers, whenever p and p_1 are rational integers of equal parity.

14. Solve the problem of 12 for the field of 13.

15. Prove that if c is an integer of the form $4k + 1$, which does not contain a perfect square factor, then the integers in the field determined by the numbers $a + b\sqrt{c}$, a and b rational, are the numbers of the form $\dfrac{p}{2} + \dfrac{p_1\sqrt{c}}{2}$, p and p_1 being rational integers of equal parity.

16. Prove that if c is an integer without square factor, which is not of the form $4k + 1$, then the integers in the field $a + b\sqrt{c}$, a and b rational, are the numbers $p + p_1\sqrt{c}$, where p and p_1 are rational integers.

73. The discovery of a new world. Let us now take up the problem of factorization in some of the fields of algebraic numbers that we have been considering. The problem is that of determining two or more integers *of the field* whose product is equal to a given integer of the same field. If we are to determine prime factors, we must know what is to be meant by "prime integers of

the field." For this purpose we carry over with suitable changes the definition of a prime number given in Definition II to the quadratic fields of algebraic numbers. We recall that 1 is a unit of every quadratic field. Leaving open the question whether there are other units in the fields we consider, we introduce the following definition:

Definition XXXIII. A *prime integer in a quadratic field* of algebraic numbers is an integer of that field different from the units and such that there are no two integers of the field, except itself and a unit, whose product is equal to the given integer.

The recognition of prime numbers is a difficult enough problem in the set of rational integers, particularly when large numbers are involved. In a field of algebraic numbers there is an additional difficulty, because for determining the factors of such numbers we do not have any tests for divisibility nor the algorithms familiar from childhood days in the field of rational integers. The task is facilitated somewhat by means of a theorem which can be looked upon as a generalization of a part of Theorem X (see p. 90). It refers to numbers of the form $a + b\sqrt{c}$, where a and b are rational numbers, and c an integer, positive or negative, which does not have a perfect square factor. Such numbers constitute a quadratic field for every fixed c. It is convenient to have a symbol for such fields; accordingly we shall use $F(\sqrt{c})$ as an abbreviation for the phrase "the field of algebraic numbers whose elements are numbers of the form $a + b\sqrt{c}$, where a and b are rational numbers and c is a fixed integer which does not contain a square factor." Notice that for any given field of this character, c is fixed while a and b roam over the field R of rational numbers; by taking different integers for c, we obtain different fields. Thus the fields considered in *71*, under (a), (b) and (c), are denoted by $F(i)$, $F(\sqrt{5})$ and $F(\sqrt{-6})$ respectively. We are now able to prepare the way for the theorem we have in mind.

Definition XXXIV. The norm of an algebraic number of the form $a + b\sqrt{c}$ belonging to the field $F(\sqrt{c})$ is the integer $a^2 - b^2c$. It is designated by the symbol $N(a + b\sqrt{c})$, so that we have

$$N(a + b\sqrt{c}) = a^2 - b^2c.$$

Applied to the complex number $a + bi$ of the field $F(i)$, this definition reduces to $a^2 + b^2$, which was defined in Definition XXI

as the norm of $a + bi$ (compare pp. 87–90); thus the use of the word "norm" and of the notation $N(a + b\sqrt{c})$ is justified. But while the norm of a complex number is always positive, the norm of an algebraic number of the field $F(\sqrt{c})$ may be positive or negative. If c is positive the field $F(\sqrt{c})$ will contain numbers whose norm is negative; but it is always positive if c is negative. It is clear moreover that in all cases

$$N(a + b\sqrt{c}) = (a + b\sqrt{c})\,(a - b\sqrt{c}).$$

It is now a simple matter to generalize that part of Theorem X which concerns the norm.

Theorem XXI. The norm of the product of two algebraic numbers belonging to the field $F(\sqrt{c})$ is equal to the product of their norms.

Proof. Let $a + b\sqrt{c}$ and $a_1 + b_1\sqrt{c}$ be two numbers of the field $F(\sqrt{c})$. Their product will be another number of the field (compare 25, 2, p. 42); let it be $r + s\sqrt{c}$; then we find easily that

$$r = aa_1 + bb_1c, \text{ and } s = ab_1 + a_1b.$$

Furthermore, it follows from Definition XXXIV, that

$$N(a + b\sqrt{c}) = a^2 - b^2c,\ N(a_1 + b_1\sqrt{c}) = a_1^2 - b_1^2c;$$

and

$$\begin{aligned} &N(r + s\sqrt{c}) = r^2 - s^2c = (aa_1 + bb_1c)^2 - (ab_1 + a_1b)^2c \\ &= a^2a_1^2 + 2aa_1bb_1c + b^2b_1^2c^2 - a^2b_1^2c - 2aa_1bb_1c - a_1^2b^2c \\ &= a^2(a_1^2 - b_1^2c) - b^2c(a_1^2 - b_1^2c) = (a^2 - b^2c)\,(a_1^2 - b_1^2c) \\ &= N(a + b\sqrt{c})\cdot N(a_1 + b_1\sqrt{c}). \end{aligned}$$

Thus our theorem is proved by the use of very simple means.

For our present purpose we are interested particularly in the norms of the integers of the field $F(\sqrt{c})$, in which c is a negative integer not of the form $4k + 1$; let us put $c = -C$, $C > 1$. If we take into account the facts stated in 72, 15 and 16, it follows that if $a + b\sqrt{-C}$ is an integer of $F(\sqrt{-C})$, then $N(a + b\sqrt{-C})$ is a rational integer of the form $p^2 + p_1^2C$, where p and p_1 are rational integers.

Theorem XXII. If the norm of an integer of the field $F(\sqrt{-C})$, where C is a positive integer greater than 1 and $-C$ is not of the

form $4k + 1$, is not the product of two rational integers of the form $p^2 + p_1^2C$ (each different from this norm itself), then this integer is a prime integer of the field.

Suppose that $a + b\sqrt{-C}$ is an integer but not a prime integer of our field $F(\sqrt{-C})$; then there exist two other integers of this field, say $m + m_1\sqrt{-C}$ and $n + n_1\sqrt{-C}$, neither of them equal to 1, such that

$$a + b\sqrt{-C} = (m + m_1\sqrt{-C})(n + n_1\sqrt{-C}).$$

But then we would have, in accordance with Theorem XXI,

$$N(a + b\sqrt{-C}) = N(m + m_1\sqrt{-C}) \cdot N(n + n_1\sqrt{-C}), \text{ i.e.}$$

$$(7.9) \qquad a^2 + b^2C = (m^2 + m_1^2C)(n^2 + n_1^2C).$$

By hypothesis $a^2 + b^2C$ does not have any factors of such form which are different from itself. Hence one of the two factors on the right side of (7.9) would have to be equal to $a^2 + b^2C$, and therefore the other factor would have to be 1. Suppose then $n^2 + n_1^2C = 1$. Since C is positive and > 1 and since n and n_1 are to be integers this relation can hold only if $n = \pm 1$ and $n_1 = 0$; i.e. if $n + n_1\sqrt{-C} = \pm 1$, so that one of the factors of $a + b\sqrt{-C}$ is a unit; this is in contradiction with our hypothesis that $a + b\sqrt{-C}$ is not a prime. Thus our theorem is proved.

Let us look at a few examples. The field $F(\sqrt{-6})$, which we have studied in *71*, (c) is of the kind covered by the last theorem. We already know that its integers are of the form $p + p_1\sqrt{-6}$ (compare p. 141) and that their norms are positive rational integers of the form $p^2 + 6p_1^2$. By means of Theorem XXII we can show that $\sqrt{-6}$ is a prime integer of the field. We find that $N(\sqrt{-6}) = 6$; the only factors of $N(\sqrt{-6})$, different from 6 and from 1, are 2 and 3, neither of which is of the form $p^2 + 6p_1^2$, as the reader will readily verify. Therefore $\sqrt{-6}$ is a prime integer of the field $F(\sqrt{-6})$.

And now we have reached the objective point of our expedition. For it is easily seen that within the field $F(\sqrt{-6})$, the integer 25 admits two sets of factors, viz. 5×5 and $(1 + 2\sqrt{-6})(1 - 2\sqrt{-6})$. Are these factors prime integers in the field? We find that $N(5) = N(1 + 2\sqrt{-6}) = N(1 - 2\sqrt{-6}) = 25$. The only rational

integral factor of 25 which is different from 25 and from 1 is the factor 5. But the number 5 can obviously not be put in the form $p^2 + 6p_1^2$; for to the values $p = 0, 1, 2$ would correspond the conditions $6p_1^2 = 5, 4, 1$ respectively, of which no one corresponds to an integral value of p_1. Hence 5, $1 + 2\sqrt{-6}$ and $1 - 2\sqrt{-6}$ are indeed prime integers of the field; and 25 is therefore factorable in at least two ways into prime integers of the field $F(\sqrt{-6})$. We conclude that *in this field the unique factorization theorem does not hold.*

We will consider next the field $F(\sqrt{-5})$; since $-5 = -6 + 1$, this field is also one to which Theorem XXII applies. Its integers are of the form $p + p_1\sqrt{-5}$, and their norm is $p^2 + 5p_1^2$. Within this field 6 admits two sets of factors, viz. $6 = 2 \cdot 3$ and $6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$. Furthermore, we can show by use of Theorem XXII that 2, 3, $1 + \sqrt{-5}$ and $1 - \sqrt{-5}$ are all prime integers of $F(\sqrt{-5})$. For $N(2) = 4$; besides 4 and 1, it has the factors 2 and 2, which are not of the form $p^2 + 5p_1^2$ for rational integral values of p and p_1. Similarly $N(3) = 9$, whose factors apart from 9 and 1 are 3 and 3, which also refuse to be put in the form $p^2 + 5p_1^2$. Finally $N(1 + \sqrt{-5}) = N(1 - \sqrt{-5}) = 6$, whose relevant factors are 2 and 3, of which we already know the attitude towards the form $p^2 + 5p_1^2$. Consequently 6 has two distinct sets of prime factors in the field $F(\sqrt{-5})$.

It is hardly necessary to give further examples of fields which do not possess the convenient property that its integers can be factored into prime factors in only one way; those we have seen may suffice to suggest that the field of rational numbers, in which most of us spend most of our lives, is perhaps exceptional. At any rate the excursion to these somewhat barbarous fields must have the effect of making us hurry home to assure ourselves that in our familiar field no such conditions exist.[1]

If our expedition has had this effect, its purpose has been accomplished. For there are a great many books which will furnish the anxious reader the assurance which he craves.

74. Everything is safe at home. There is little point in reproducing here a proof of the fact that the rational integers can be

[1] In *20*, 7 examples are found of sets of natural numbers, which are not fields, but in which unique factorization fails.

factored into prime factors in only one way. All we are concerned with is the recognition of the need for such a proof. There are many good guides for those who wish to undertake this examination of the foundations upon which so much of our work with integers depends. Indeed, every book which deals with the theory of numbers contains a proof of the theorem. The reader will find one on page 3 of L. E. Dickson's *Introduction to the Theory of Numbers;* also on pages 15 and 16 of R. D. Carmichael's *The Theory of Numbers.* A very instructive discussion of the subject matter of the present chapter is found in the charming book called *Von Zahlen und Figuren* by H. Rademacher and O. Toeplitz; an excellent treatment occurs in J. Sommer, *Introduction à la théorie des nombres algébriques.*

A few exercises on algebraic numbers will help the reader to acquire a mastery over the road we have just traveled.

75. To become acquainted with the new world.

1. Show that $\sqrt{-5}$ and $4+\sqrt{-5}$ are prime integers of the field $F(\sqrt{-5})$.

2. Show that if c is a positive integer, not of the form $4k+1$, which does not have a square factor, then $\sqrt{-c}$ is a prime integer of the field $F(\sqrt{-c})$.

3. Determine factors of $\sqrt{6}$ in the field $F(\sqrt{6})$.

Hint. Theorem XXII is not very useful for the determination of factors of non-prime integers — its value has been a negative one. A direct attack on the problem can be made by putting $\sqrt{6} = (a+b\sqrt{6})(c+d\sqrt{6})$ and seeking rational integers a, b, c and d such that $0 = ac + 6bd$, and $1 = ad + bc$.

4. Prove that if α is an integer of the field $F(\sqrt{-7})$ then $N(\alpha)$ is a positive rational integer.

5. Prove that the norm of an integer of any field $F(\sqrt{c})$ is always a rational integer.

6. Show that in the field $F(\sqrt{-7})$, the integer 2 has the integral factors $\frac{1}{2}+\frac{\sqrt{-7}}{2}$ and $\frac{1}{2}-\frac{\sqrt{-7}}{2}$; are these prime integers of the field?

7. Prove that an integer α of the field $F(\sqrt{-7})$ is a prime integer, if $N(\alpha)$ does not have two factors of the form n^2+7m^2 or of the form $n^2+7m^2+n+7m+2$ and each distinct from $N(\alpha)$.

8. Prove that the norm of a unit in a quadratic field is equal to $+1$ or to -1.

9. Prove that if the norm of an integer in a quadratic field is equal to $+1$ or -1, then this integer is a unit of the field.

10. Prove that Theorem XXII holds also for fields $F(\sqrt{C})$, if C is a positive integer, without square factor and not of the form $4k+1$.

11. Determine factors in $F(\sqrt{5})$ of $1+\sqrt{5}$.

12. Show that in $F(\sqrt{-6})$, the integer 6 admits two distinct sets of prime factors.

13. Prove that if $\frac{p}{2}+\frac{p_1\sqrt{-7}}{2}$ and $\frac{s}{2}+\frac{s_1\sqrt{-7}}{2}$ are integers of the field $F(\sqrt{-7})$, then their quotient is also an integer of the field if and only if $2(ps+7p_1s_1)$ and $2(ps_1-p_1s)$ are divisible in the set of rational integers, by $s^2+7s_1^2$, and provided the quotients are of equal parity.

14. Determine the units of the field $F(i)$.

15. Prove that Theorem XXII still holds if the condition $C>1$ is omitted.

76. The anatomy of truth. The failure of the unique factorization theorem in the fields $F(\sqrt{-5})$ and $F(\sqrt{-6})$ carries with it the non-validity of conclusions in the derivation of which this theorem plays an essential rôle, that is to say of conclusions which could not be reached without its use. Thus a thorough study of fields like $F(\sqrt{-5})$ and $F(\sqrt{-6})$ will have as one of its advantages insight into the real significance of the unique factorization theorem. It enables us to divide familiar results into two groups, one containing those which depend on unique factorization and the other such as do not depend upon unique factorization. In general we gain insight into the structure of a body of "truth," by the study of a world in which the presuppositions are in contrast with those of the familiar world — the study of a foreign civilization derives much of its real value from the fact that it sets up a picture contrasting with the domestic one. It is a not unimportant aspect of our study of algebraic numbers that it leads to an appreciation of such values.

Returning to the study of $F(\sqrt{-5})$ we can show that unique factorization is essential for Theorem I (compare p. 10). According to this theorem, *every* primitive Pythagorean triple is obtainable from (1.2), if u and v are relatively prime integers in that field. Now we have already seen on page 147, that 2 and $1-\sqrt{-5}$ are

prime integers of $F(\sqrt{-5})$; and in *75*, 1, it was shown that $4+\sqrt{-5}$ and $\sqrt{-5}$ are also prime integers of this field. Consequently, the integers $2-2\sqrt{-5}$, $4+\sqrt{-5}$ and $\sqrt{-5}$ do not have any factor in common; moreover

$$(2-2\sqrt{-5})^2+(4+\sqrt{-5})^2=(\sqrt{-5})^2,$$

so that $(2-2\sqrt{-5},\ 4+\sqrt{-5};\ \sqrt{-5})$ is a primitive Pythagorean triple in $F(\sqrt{-5})$. But it is not of the form given in (1.2), since neither $2-2\sqrt{-5}$ nor $4+\sqrt{-5}$ can be effectively represented as $2uv$. This is at once evident for $4+\sqrt{-5}$, since it is a prime integer of the field. That it is also true for $2-2\sqrt{-5}$ follows from the fact that the only factors of $1-\sqrt{-5}$ are 1 and $1-\sqrt{-5}$; but $u=1$, and $v=1-\sqrt{-5}$ leads to $u^2+v^2=-3-2\sqrt{-5}$ and not to $\sqrt{-5}$. Therefore while (1.2) gives primitive Pythagorean triples also in the field $F(\sqrt{-5})$, it does not furnish all of them; for there are primitive Pythagorean triples in this field which can not be obtained from (1.2); thus Theorem I fails.

It would be worth while to illustrate in other ways the consequences of the non-validity of the unique factorization theorem. But we shall not pursue this trail any further. Our aim has been accomplished with the realization that there are other worlds in the domain of mathematics besides that of rational numbers with which ordinary experience is primarily concerned, that in some of these worlds the fundamental constitution is different from that of the world with which we are familiar, and that these differences may lead to a new orientation. There is here again a tempting parallel with concrete human situations. We are too prone to assume that in every human group, be it a nation, or a city, or any of the numerous associations of human beings, the fundamental bases must be the same as those for the corresponding group in which we find ourselves, and thus to attribute differences in conduct to moral depravity, or perhaps to uncommon excellence. It would be well to inquire before making such a conclusion whether such identity of basal principles actually exists. In judging the Chinese people, or the Mexican Indians, or people of a different social heritage in our midst, it is well to remember if we do not want to condemn unfairly or to praise unduly, that different groups

may have built on foundations quite different from one another. Just as in fields of algebraic numbers conditions exist which are incomprehensible to one who has never been outside the field of rational numbers, but which follow inevitably from the fundamental characteristics of such fields, so human groups different from the one with which we are acquainted may quite naturally adhere to customs and convictions which appear to us as contrary to "human nature."

77. The creative power of failure. We have seen repeatedly that the failure of an operation within a given set of numbers has given rise to a fruitful enlargement of this set. The desire to remove exceptions to the validity of a theorem or to the possibility of a procedure has led to the creation of new entities. For instance, when we found that subtraction was not possible without exception in the set of natural numbers, we established the set of all integers, positive, negative or zero, a part of which is isomorphic with the set of natural numbers with respect to addition and multiplication, and within which subtraction is possible without exceptions. The desire to make the solution of the equation $x^n = a$ possible for every rational number a and every rational n led to the introduction of the real numbers and the complex numbers. It is quite to be expected therefore that mathematicians were spurred on by the failure of the unique factorization theorem in certain fields of algebraic numbers to the search for new entities which would restore this theorem. This has actually been done by the work of a number of men who, in the course of the 19th century, created the *theory of ideals*. Among the large number of those who have contributed to the development of this theory we shall only mention Eduard Kummer (1810–1893), P. G. Lejeune Dirichlet (1805–1859) and Richard Dedekind (1831–1916), whom we have met on an earlier occasion (compare pp. 60, 63). In any of the large number of books devoted to the theory of algebraic numbers ample references to other writers will be found. It would be a fitting close to our excursion in the fields of algebraic numbers to explore the theory of ideals. For it is one of the most beautiful mathematical creations of the 19th century, and it plays a rôle of great significance in much work that is being done to-day. But such an undertaking would not fit into our present plans. Let it therefore be reserved for a later time and for those readers who are prepared for the concentrated attention which it requires.

78. Squares and cubes. The greater part of this chapter has been concerned with a few special quadratic fields, that is to say with sets of numbers which satisfy certain equations of type (7.3), which is itself but a special case of the general equation (7.2). The reason for this limitation is the fact that the reader while probably familiar with the roots (7.4) of a quadratic equation is not likely to have had any experience with equations of type (7.2) in which n exceeds 2. It will be worth while for us to give some attention to such equations now. For $n = 3$, we obtain from (7.2) the *general cubic equation*

$$a_0x^3 + a_1x^2 + a_2x + a_3 = 0. \tag{7.10}$$

In the study of algebra, methods are developed for determining complex numbers which satisfy this equation. One also finds a systematic treatment of the case $n = 4$, which deals with the *general biquadratic equation.*

$$a_0x^4 + a_1x^3 + a_2x^2 + a_3x + a_4 = 0. \tag{7.11}$$

But, in accordance with our general point of view, we shall not be concerned with these methods. It has already been *stated* (compare p. 134) that the set of complex numbers contains all numbers which satisfy any equation of either of these types. The roots may be rational numbers, or quadratic irrational numbers; in such cases they also satisfy equations of type (7.2) of degree lower than 3 and 4. For instance, it is immediately verifiable that 2, $1 + \sqrt{-3}$ and $1 - \sqrt{-3}$ satisfy the equation $x^3 - 4x^2 + 8x - 8 = 0$, which is of type (7.10), but 2 also satisfies the equation of the 1st degree with integral coefficients $x - 2 = 0$, while $1 + \sqrt{-3}$ and $1 - \sqrt{-3}$ are also roots of the quadratic equation $x^2 - 2x + 4 = 0$, which is of type (7.3). Again, $2 + \sqrt{5}$, $2 - \sqrt{5}$, $3 + \sqrt{2}$ and $3 - \sqrt{2}$ satisfy the equation $x^4 - 10x^3 + 30x^2 - 22x - 7 = 0$, which is of type (7.11); on the other hand $2 + \sqrt{5}$ and $2 - \sqrt{5}$ belong to the quadratic equation $x^2 - 4x - 1 = 0$, and $3 + \sqrt{2}$, $3 - \sqrt{2}$ to the quadratic equation $x^2 - 6x + 7 = 0$. There are however many equations of the 3rd and 4th degrees which are not satisfied by numbers that belong to equations of lower degree. Their roots come from parts of the set of algebraic numbers differ-

ent from the rational and quadratic fields. Such equations are called *irreducible equations.*[1]

Any number which satisfies an irreducible cubic equation, i.e. a number which satisfies an equation of type (7.10) but not any equation of degree lower than 3, is called a *cubic irrational;* it belongs to a cubic field. In the same way we have biquadratic fields, quintic fields and so forth. An example of a cubic irrational is $\sqrt[3]{2}$, which satisfies the equation $x^3 - 2 = 0$, but not any equation of type (7.2) of degree lower than 3; the equation $x^3 - 2 = 0$ is an irreducible cubic equation.

With this last equation, a famous problem of antiquity is closely associated. On pages 135 and 136, we were led to pay some attention to the problem of squaring the circle. A second problem, with which the Greeks occupied themselves a great deal, is that of the *duplication of the cube;* it requires the construction of a cube twice as large in volume as a given cube. In the discussion of the circle-squaring it was observed that to require a construction without specifying the tools that are available is rather meaningless. As with the former problem so with the duplication of the cube, it is often not recognized that the question takes on special significance and interest when we specify that the construction is to be carried out by the use of a straight edge and a pair of compasses only.[2]

A history of the problem in antiquity is found on pages 154–170 of Heath's *Manual of Greek Mathematics*, to which we have referred on many earlier occasions; a discussion from the modern point of view can be read in Klein's *Famous Problems* (compare p. 136). There it is shown that the construction can not be carried out by merely using a straight edge and a pair of compasses. The proof of the impossibility of the construction is much simpler in this case than in the case of the squaring of the circle. It is based on the following theorems: (a) "The necessary and sufficient condition that an analytic expression can be constructed with straight edge and compasses is that it can be derived from the known quantities by a finite number of rational operations and square roots"[3]; and (b) "If an irreducible equation is not of degree 2^h, it can not be solved by square roots."[4]

[1] More accurately, "irreducible in the field R of rational numbers."
[2] Compare J. W. A. Young, *Monographs on Topics of Modern Mathematics*, p. 353.
[3] See Klein, *op. cit.*, p. 3; Young, *op. cit.*, p. 354.
[4] See Klein, *op. cit.*, p. 12; Young, *op. cit.*, p. 363.

Since the equation $x^3 - 2 = 0$ is not reducible and since its degree is 3, which does not have the form 2^h, the root $\sqrt[3]{2}$ can not be constructed by straight edge and compasses. But the problem of duplicating the cube leads exactly to the equation $x^3 - 2 = 0$. For if we use the side of the given cube as the unit of length and denote the length of a side of the required cube by x, the volumes of the given and the required cubes will be equal respectively to 1 and x^3 volume units. Hence the problem requires the "construction" of a line of length x, such that $x^3 = 2$. The impossibility of this construction is therefore an immediate consequence of the theorems quoted under (a) and (b).

While we are in this vicinity we may well have a look at the third famous problem of antiquity, that of the "trisection of an arbitrary angle." These three problems, "the squaring of the circle," "the duplication of the cube," and "the trisection of an arbitrary angle" have become celebrated for the unremitting attempts at solving them made throughout the ages. It is a curious fact that all three have in modern times been proved incapable of solution, it being understood that in each case only straight edge and compasses are to be employed. The last mentioned problem, that of the trisection of an arbitrary angle, continues to attract attention perhaps even more than the others, in spite of the fact that a proof of its impossibility is within the reach of every college student of mathematics. It is probably no exaggeration to say that every year the mathematics departments of colleges and universities throughout this country receive numerous "solutions" of this problem. They come from bakers and butchers and candlestick makers. Not infrequently a high school student succeeds in making the front page of a local newspaper with the announcement of his discovery. It is a sad commentary on the teaching of mathematics in the schools that interested and ambitious pupils are allowed to waste their energies on so futile an undertaking instead of being directed into more fruitful fields where their time can be employed in a valuable way. But let us return to the problem of the trisection of the angle.

79. The trisection of the angle. Let us recall some of our earlier experiences, and let us agree that *construct* is to mean *construct with straight edge and compasses.* It should be clear from Definition XXII and Fig. 23 (see p. 98), that if an angle θ can be constructed, then a line equal to $\cos \theta$ can also be constructed. For once the

angle θ has been found, we have but to construct an angle equal to θ in *standard position*, lay off a unit distance on its terminal side, from the origin to a point P, and then drop a perpendicular from P to the X-axis. All these constructions can be carried out by means of straight edge and compasses; they are familiar to every one who has studied plane geometry. If the foot of the perpendicular is R, then $OR = \cos \theta$. Conversely, if a line equal to $\cos \theta$ can be constructed then θ can be found by laying off on the X-axis a line OR equal to $\cos \theta$, erecting a perpendicular to the X-axis at R and finding the intersection P of this perpendicular with the unit-circle about O; then $\angle XOP = \theta$. We have therefore established the following result:

Theorem XXIII. The construction of an angle θ and the construction of a line equal to $\cos \theta$ are either both possible or both impossible.

The reader who has carried out *56*, 14 (p. 101) will know that

$$\cos 3\theta = 4(\cos \theta)^3 - 3 \cos \theta \tag{7.12}$$

for every value of θ. (If he has not carried it out, he will have to take (7.12) on the authority of those who have.)

The problem of trisecting an arbitrary angle can be put in the following form, by the use of Theorem XXIII: Given a line a equal to $\cos 3\theta$, to construct a line x equal to $\cos \theta$; or again, by use of (7.12): Given a line a to construct a line x such that

$$a = 4x^3 - 3x. \tag{7.13}$$

Having reached this formulation of the problem it is not difficult to show that not every angle can be trisected. For, if $3\theta = 120°$, $a = \cos 3\theta = -\frac{1}{2}$ (see p. 99); equation (7.13) becomes in this case $-\frac{1}{2} = 4x^3 - 3x$, or $8x^3 - 6x + 1 = 0$. It can be reduced to a more convenient form by putting $x = \frac{y}{2}$; for then it becomes simply

$$y^3 - 3y + 1 = 0. \tag{7.14}$$

If we can show that this equation is irreducible, then we can conclude by means of (a) and (b) on page 153, as in the case of the duplication of the cube, that no root of (7.14) can be constructed, which means that our problem is not capable of solution.

To prove that (7.14) is indeed irreducible, we convince ourselves

first that no rational number satisfies it. Suppose that $\frac{p}{q}$ were a root of (7.14), p and q being relatively prime integers. It would then follow that $\frac{p^3}{q^3} - \frac{3p}{q} + 1 = 0$, or that

$$p^3 = (3p - q)q^2. \tag{7.15}$$

Since the right side of this identity contains the factor q^2, the unique factorization theorem (it holds in the field of rational integers!) tells us that p^3 must contain this factor as well. On the other hand p and q are relatively prime integers. These two conditions are compatible only if $q = 1$. Substituting 1 for q in (7.15), this relation becomes $p^3 = 3p - 1$, or $p(p^2 - 3) = -1$. And now it is readily seen that there is no integer p for which this condition is satisfied. For it follows again from the unique factorization of rational integers, that we have to have either $p = 1$ and $p^2 - 3 = -1$, which are incompatible conditions, or $p = -1$ and $p^2 - 3 = +1$, which are also irreconcilable. Consequently, *equation* (7.14) *can have no rational root.*

Suppose next that (7.14) had a quadratic irrational as one of its roots, i.e. an algebraic number of the form $a + b\sqrt{c}$, where a and b are rational numbers, $b \neq 0$, and c an integer, $\neq 1$, positive or negative, without square factor (compare p. 144). It would then follow that $(a + b\sqrt{c})^3 - 3(a + b\sqrt{c}) + 1 = 0$, i.e. that

$$a^3 + 3ab^2c - 3a + 1 + (3a^2b + b^3c - 3b)\sqrt{c} = 0.$$

This relation can hold only if

$$a^3 + 3ab^2c - 3a + 1 = 0 \text{ and} \tag{7.151}$$
$$3a^2b + b^3c - 3b = 0; \tag{7.152}$$

for otherwise $\sqrt{c}$ would be equal to the negative of the rational number obtained by dividing the left side of (7.151) by that of (7.152).

From (7.152) would follow, since $b \neq 0$, that $3a^2 + b^2c - 3 = 0$, or that $b^2c = 3 - 3a^2$. Substitution of this value for b^2c in (7.151) reduces it to the form $-8a^3 + 6a + 1 = 0$. If we put $a = \frac{z}{2}$, we obtain the condition that the equation $z^3 - 3z - 1 = 0$ must be satisfied by a rational number. This equation is almost the same

as (7.14); the treatment accorded to the latter equation shows, if applied to $z^3 - 3z - 1 = 0$, that *it* does not have a rational root. Therefore no number of the form $a + b\sqrt{c}$ can satisfy (7.14). We have now completed the proof of the irreducibility of this equation, and hence by the argument of page 155, that the angle of 120° can not be trisected by straight edge and compasses. Remembering the power of a single example in throwing out a general rule (compare pp. 49, 51, 102) we have proved the following theorem:

Theorem XXIV. The trisection of an arbitrary angle by straight edge and compasses alone is impossible.

It may not be superfluous to remark that this theorem does *not* say that *no* angle can be trisected with these tools. In fact we already know that an angle of 180° can be so trisected; for, by constructing an equilateral triangle, we obtain an angle of 60°, which is equal to one third of 180°.

80. Possible constructions. The proofs of theorems (a) and (b) quoted on page 153 would carry us beyond our depth; we had therefore better not undertake them. It is however not difficult to see that any expression which "can be derived from the known quantities by a finite number of rational operations and square roots" is indeed constructible. As far as the rational operations are concerned, we have already dealt with them in Chapter V (compare pp. 72–75); for, the rational operations on points which we used there can at once be interpreted as operations on the lines which join the origin to these points. It only remains to show that a square root of a known quantity is constructible.

The known quantity can always be represented in the form ab, where both a and b are known quantities (if the known quantity is measured by a prime integer, we can still represent it in that way; for instance, we can put $7 = 7 \times 1$, or $7 = 5 \times \frac{7}{5}$, etc.). We have therefore to consider the construction of a line x such that $x^2 = ab$. The reader may remember from his school days how this is done; but we will suppose that he does not remember it. If we lay off on a straight line, segments equal to a and to b, in opposite directions from a common starting point P (see Fig. 24), describe a semi-circle on AB as a diameter, erect a perpendicular to AB at P, which meets the semi-circle at Q, and then connect Q with A and B, we obtain two right-angled triangles APQ and QPB which are similar, because the angles are equal as designated. (The reader will have no difficulty in checking up the fact that all these constructions

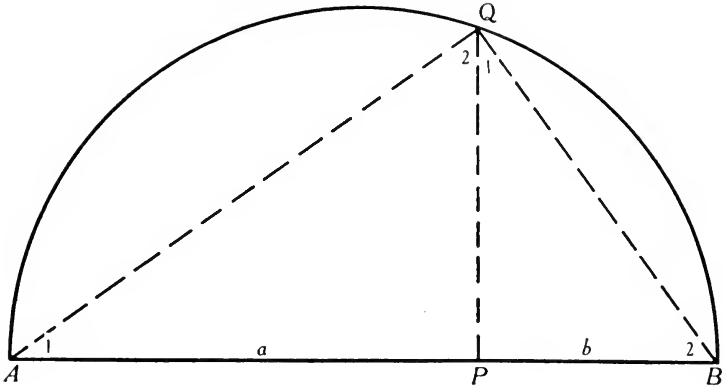


FIG. 24

can be carried out with straight edge and compasses, nor in verifying the indicated equalities of the angles, if he remembers that $\angle AQB$, inscribed in a semi-circle, is a right angle.) Therefore the sides opposite the equal angles are in proportion:

$$QP : PB = AP : QP, \text{ i.e. } \overline{QP}^2 = AP \cdot PB = ab.$$

Hence the line PQ solves the problem.

A second method for constructing a line equal to $\sqrt{ab}$ is closely connected with a proof of the theorem of Pythagoras, which was the starting point of our wanderings. We lay off on a straight line segments equal to a and b, but now in the same direction from a common starting point P (see Fig. 25). Supposing that a is the longer of the two given lines, we erect at B a perpendicular to PA, meeting in Q the semi-circle constructed on PA as a diameter. Connecting Q with P and with A, we obtain the two similar right angled triangles PBQ and PQA. The proportionality of corresponding sides now leads to the relation $PQ : PB = PA : PQ$, so that $\overline{PQ}^2 = PB \cdot PA = ab$; the line PQ in Fig. 25 is therefore also a solution of the equation $x^2 = ab$.[1]

We have therefore shown that any expression which involves

[1] By observing that the triangles AQB and APQ are also similar, we obtain, in addition to the relation $\overline{PQ}^2 = PB \cdot PA$ proved above, the further equality $\overline{AQ}^2 = AB \cdot AP$. If we add the two, we find that $\overline{PQ}^2 + \overline{AQ}^2 = PB \cdot PA + AB \cdot AP = PA(PB + BA) = \overline{PA}^2$, which proves the theorem of Pythagoras.

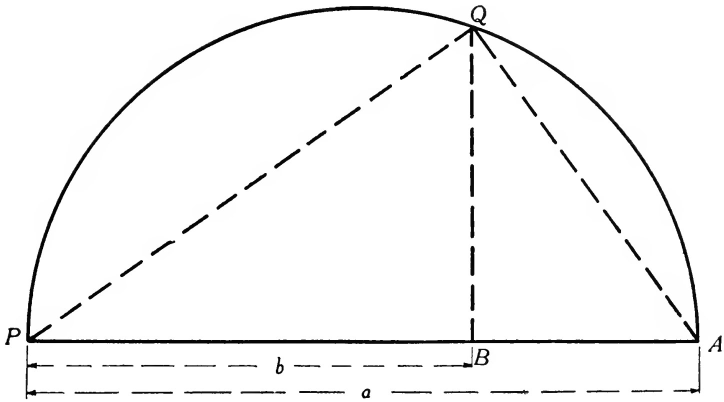


FIG. 25

only rational operations and square root extractions to be performed upon known quantities, can be constructed by means of straight edge and compasses; in other words that the condition of Theorem (a) on page 153 is sufficient for the constructibility of a segment.

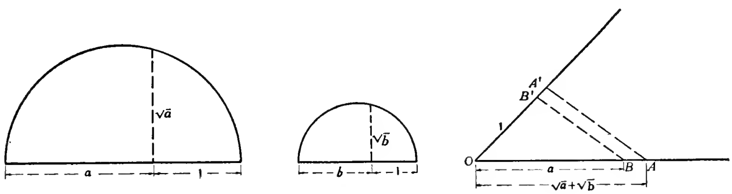


FIG. 26

Example. To construct from two given segments a and b a line equal to $\dfrac{\sqrt{a}+\sqrt{b}}{a}$: We use the construction of Fig. 24 for $\sqrt{a}$ and $\sqrt{b}$ (see Fig. 26); after that we carry out the construction of Fig. 11 (see p. 75) on the segments $\sqrt{a}+\sqrt{b}$ and a. In this way we obtain the segment OA' equal to $\dfrac{\sqrt{a}+\sqrt{b}}{a}$.

81. To secure the position.

1. Determine the equation of type (7.2) which is satisfied by $\sqrt{2}+\sqrt{6}$.

2. Determine the equation of type (7.2) which is satisfied by $\sqrt[3]{2} - \sqrt{2}$.

3. Show that if $5 - \sqrt{-3}$ satisfies a cubic equation of type (7.10), then $5 + \sqrt{-3}$ satisfies the same equation.

4. Prove that no rational number satisfies the equation $z^3 - 3z - 1 = 0$ (compare p. 157).

5. Prove that the equation $z^3 - 3z - 1 = 0$ is irreducible.

6. Prove that the construction of an angle θ and the construction of a line equal to $\sin\theta$ are either both possible or both impossible.

7. Prove that the equation $x^3 + x^2 - 2x - 1 = 0$ has no rational root.

8. Prove that the equation $x^3 + 2x^2 - x - 1 = 0$ has no rational root.

9. Prove that the equation $x^3 + x^2 - 2x - 1 = 0$ is irreducible.

10. Construct by means of straight edge and compasses a line equal to $\dfrac{\sqrt{ab} + \sqrt{a}}{\sqrt{a} - 1}$, the lines a and b being given.

11. Construct a line equal to $\sqrt{\dfrac{a}{b} + ab}$.

12. Construct a line equal to $\sqrt{\sqrt{a} + \sqrt{b}}$.

82. The preponderance of transcendental numbers. When we met the transcendental (i.e. non-algebraic) real numbers on page 137, the statement was made that they were more numerous than the algebraic real numbers. The proof of this statement is the last goal of our present excursion. Our purpose will be more readily accomplished if we prove first a theorem which completes in an interesting way some of the discoveries that turned up in Chapter II.

Theorem XXV. The set which is obtained by putting into one set all the elements of a denumerable collection of denumerable sets is itself denumerable.

Proof. To prove this theorem we have merely to reinterpret the method and diagram used in the proof of Theorem II (compare pp. 22 and 23). Let the dots in the lowest horizontal line in Fig. 4 represent the elements of the first denumerable set, those in the next horizontal line the elements of the second denumerable set, and so on. The elements of the set which we wish to prove denumerable will then be represented by the points in the plane with positive integral coördinates. But this set was proved to be denumerable in Theorem II; our present theorem is therefore proved.

Corollary 1. The set which is obtained by putting into one set all the elements of a finite number of denumerable sets is itself denumerable.

This corollary can either be proved by combining *17*, 13 and Theorem XXV, or it can be proved directly; the details are left to the reader.

It is a consequence of this corollary that the set of all integers, positive, zero or negative, is denumerable, for it consists of the single element 0, the denumerable set of positive integers and the denumerable set of negative integers.

Corollary 2. The set of all integers is denumerable.

Let us now consider the set of equations of type (7.1), which we shall write in the form $a_0x + a_1 = 0$, in which a_0 and a_1 are integers and $a_0 \neq 0$, so as to exhibit it as a special case of (7.2). We shall denote by A_1 the sum of a_0 and a_1, $A_1 = a_0 + a_1$; it is an integer and we shall call it the *height* of the equation (7.1). The total set of equations of type (7.1) can then be thought of as obtained by putting into one set all the equations for which the height A_1 is equal to 0, or 1, or 2 etc. . . ., or $-$ 1, or $-$ 2, etc. . . . This total set consists therefore, in virtue of Corollary 2, of a denumerable collection of sets of equations. What about the set of equations of fixed height? If A_1 is given, a_0 varies over the set of integers (excepting 0); for each value of a_0, there is one value for a_1, viz. $A_1 - a_0$ and hence one equation. Therefore the set of equations of any given height is denumerable. The total set of equations of type (7.1) has thus been shown to be a *denumerable* collection of denumerable sets of equations; hence, by Theorem XXV, it is itself denumerable. Since, moreover, each such equation has just one root (compare p. 132), we conclude that the roots of the equations of type (7.1), viz. the rational numbers, constitute a denumerable set.

The argument which we have just gone through may appear unnecessarily complicated, particularly since the final conclusion, i.e. the denumerability of the set of rational numbers, was already known from Chapter II. The reader will do well however to think it through carefully point by point, because it serves as a basis and model for the further discussion.

We come next to the irreducible quadratic equations of type (7.3). We define the height to be $A_2 = a_0 + a_1 + a_2$. The total set consists then of a denumerable set of sets of equations, namely

of those sets for which A_2 is equal to the different integers. What about the equations of given height? As before a_0 ranges over the set of integers, excepting 0; if a_0 is any fixed integer a then $a_1 + a_2$ is fixed at $A_2 - a$. But, by the argument made in the second paragraph above, we conclude that, if $a_1 + a_2$ is fixed, there is a denumerable set of pairs of values for (a_1, a_2). The set of values for the triple (a_0, a_1, a_2) for any given height A_2 is therefore a denumerable set of denumerable sets, hence itself denumerable by Theorem XXV. Applying the same theorem again we conclude that the set of equations of type (7.3) is denumerable. The subset of irreducible quadratic equations is then certainly denumerable (compare *17*, 13). Every such equation has two roots (compare p. 133); the roots of all these equations can then be grouped into two denumerable sets and constitute themselves a denumerable set, on the strength of Corollary 1. Therefore the set of quadratic irrationals constitutes a denumerable set (compare *72*, 3).

Stepping up to the case $n = 3$ we come to the cubic equations, those of type (7.10). The height is $A_3 = a_0 + a_1 + a_2 + a_3$. For a given height and a given fixed value a of a_0 ($\neq 0$), we have $a_1 + a_2 + a_3$ fixed at $A_3 - a$. For this fixed value, the set of values for the triple (a_1, a_2, a_3) is seen to be denumerable by a mere repetition of the argument of the preceding paragraph. The set of tetrads (a_0, a_1, a_2, a_3) is therefore denumerable for a fixed height A_3; and, as A_3 runs through all its possible values, the set of equations of type (7.10) remains denumerable, thanks once more to Theorem XXV.

To carry on the argument from this point, we need an important fact which follows from the fundamental theorem of algebra (compare p. 134) and which we shall state without proof, viz. the following:

(7.16) "Every equation of type (7.2) has exactly n roots in the field of complex numbers."

In particular the cubic equation has exactly 3 roots. The set of roots of the irreducible equations of type (7.10) can then be grouped into 3 denumerable sets and is therefore denumerable.

In this way we proceed step by step. After we have proved that the set of values of $a_0, a_1, \ldots a_n$ for which $a_0 + a_1 + \cdots + a_n$ is fixed, is denumerable, it follows that the set of values of $a_0, a_1, \ldots a_n, a_{n+1}$ for which $a_0 + a_1 + \cdots + a_n + a_{n+1}$ is fixed consists of a denumerable set of denumerable sets and is therefore

itself denumerable. Then by the same argument that we have now used several times we conclude that the set of roots of the irreducible equations of type (7.2) is denumerable for every *fixed* positive integral value of n.[1]

Now comes the final step. The possible degrees of equations of type (7.2) form a denumerable set. The totality of all the roots of *all* equations of this type form therefore a denumerable collection of denumerable sets and hence a denumerable set. We are now justified in stating the following striking theorem:

Theorem XXVI. The set of all algebraic numbers is denumerable.

By applying *17*, 13 once more, it follows (since not all algebraic numbers are real!) that the set of all real algebraic numbers is denumerable.

Having brought one more set of numbers into the fold of denumerable sets, it is well to recall that the set of all real numbers is not denumerable (compare Theorem IV, p. 29, and *39*, 9). Since every real number which is not an algebraic number is a transcendental real number, it follows from our last remark in conjunction with Theorem XXVI that the set of transcendental real numbers is not denumerable. For if it were, we would be led by use of Corollary 1 of Theorem XXV to the conclusion that the set of real numbers is denumerable; but this would contradict Theorem IV. We record this last conclusion:

Theorem XXVII. The set of transcendental real numbers is not denumerable.

In Theorems XXVI and XXVII we have an exact statement of what was vaguely foreshadowed at the opening of this section and on page 137.

It can not fail to seem very strange to the wanderer through these domains that, in spite of such wide prevalence of transcendental numbers, he has actually been introduced to only two of them on all his journeys up to date, viz. π and e (compare p. 137); perhaps he has seen a few more in *72*, 9 and 10. Where do they keep themselves?

The search for transcendental numbers is one of the pleasures which await those who are willing to go in for the further study of mathematics. Although we shall probably get several more

[1] Notice that the principle of mathematical induction is here once more brought into play (compare p. 91).

glimpses of them on our future wanderings, our equipment will not suffice to bring them out of their hiding places into the clear light of understanding. By means of the important theorem quoted on page 136, the reader will be able to ferret out a considerable lot of them. We have had a striking illustration of the way in which mathematics has through persistent abstraction and generalization distilled out of the raw material of concrete human experience gems of rare beauty and ideas of profound significance. We have had to journey a long way from the rational numbers of everyday life to reach these wider domains. Still greater are the treasures of beauty and of understanding which reveal themselves if we carry further and further the processes of abstraction and of generalization.

CHAPTER VIII

A RETURN TO THE SIMPLE LIFE

The long, slow process, old as the race, through which the frontiers of the known have steadily encroached upon the territory of the unexplored, has been a progressive conquest of new worlds for the imagination. — John L. Lowes, *The Road to Xanadu*, p. 113.

83. Adventure within the walls. The excursions of the last several chapters have acquainted us with the domain of complex numbers; there we found greater freedom to operate than was possible in the set of rational numbers. Indeed, in this wider domain we could carry out without restriction the operations of addition, multiplication, and their inverses, subtraction and division (excepting always division by zero!); the operation of involution, with an arbitrary complex number as exponent, and its two inverses, the extraction of roots and the determination of logarithms. With respect to all these operations, the set of complex numbers is *closed*.

After having followed some of the main highways in this region we traced out some crossroads; in this way we were able to mark off a territory in which customs and manners reigned of which our earlier sheltered existence had given no inkling. Those who stay at home run fewer risks, but they also miss the exhilaration which comes to the more intrepid spirits who venture forth!

It would however be an unfortunate effect of these journeys if they had completely spoiled for us the enjoyment of the simple life. Indeed it is hoped that on the contrary they have made us remember the peculiar charms of the domain of the natural numbers; perhaps the saying of Kronecker quoted on page 36 will be recalled. If so, we should be ready now to undertake visits to some points of interest near our starting point; let us then return to the scene of the first excursion.

84. The legacy of a great man. We began by looking at and around the equation

$$(1.1) \qquad a^2 + b^2 = c^2;$$

as a result of our examination, we found that there are many triples of natural numbers, they were called Pythagorean triples, which could be substituted for a, b, c, so as to verify equation (1.1). Indeed we concluded that if u and v are any two natural numbers, relatively prime and of different parity, $u > v$, then $a = u^2 - v^2$, $b = 2uv$, $c = u^2 + v^2$ always satisfy this equation. Moreover, they will be relatively prime — they constitute a primitive Pythagorean triple; of such triples there is a denumerable infinity (compare *17*, 7) and all of them are represented by these formulas.

What then is more natural than to ask for solutions of the equations $a^3 + b^3 = c^3$, or $a^4 + b^4 = c^4$, or $a^5 + b^5 = c^5$, and so forth? It must always be understood that by a solution of such equations we shall mean triples of *natural* numbers, which may be substituted for a, b and c so as to verify these equations.

The reader is by now sufficiently familiar with mathematical methods to expect that the entire denumerable set of equations that now lies before us will be considered as a whole, as far as possible. We represent them in the form

$$(8.1) \qquad x^n + y^n = z^n,\ n > 2,$$

it being understood that n is a natural number and that we seek natural numbers for x, y and z.

Equation (8.1) has an interesting history. In the second book of the *Arithmetica*, by the Greek mathematician Diophantus of Alexandria, who lived during the latter part of the 3rd and the first decades of the 2nd centuries,[1] the problem with which we were concerned in the first chapter appears in the following form: "To divide a square number into two other square numbers." A Latin-Greek edition of the *Arithmetica* was published by C. G. Bachet, Seigneur de Meziriac, in 1621. A later reprint of this edition aroused the interest of Pierre de Fermat (1601–1665), Counselor at the parliament of Toulouse. Although mathematics was for him an occupation of leisure hours, Fermat has made many important contributions to various fields of mathematics. In the margin of his copy of Bachet's edition of Diophantus, opposite the Pythagorean problem, Fermat wrote (in Latin): *But, to divide a cube into two cubes, a fourth power into two fourth powers and in general*

[1] Not much seems to be known about the life of Diophantus; most of the information concerning him is found in T. L. Heath's *Diophantus of Alexandria*. Compare also books on the history of mathematics.

any power whatever, above the second, into two powers of the same denomination is impossible, of which fact I have indeed discovered a marvelous proof. This margin is too narrow to hold it.[1] This legacy of Fermat's has been a very fertile one. For no indications of his "marvelous proof" have ever been discovered. And, in spite of much work by many brilliant mathematicians during the last 75 years, it has not been possible either to prove or to disprove his statement. It is frequently referred to as "Fermat's last problem" or as "the great Fermat theorem"; some one has called it a challenge to human insight. Here we have then an unsolved problem concerning natural numbers, the simplest domain in mathematics; its statement is moreover easily comprehensible to every one.

To disprove Fermat's theorem requires but a single example of three integers such that a power above the second of one of them is equal to the sum of the same powers of the other two. To prove it a good deal more is required. The very simple form in which the problem is stated has tempted a great many tailors and butchers and candlestick makers, particularly since a prize award was established for its solution. Nevertheless, it still awaits its Columbus.

A good deal of progress has been made on it in recent years. More important perhaps than this progress is the stimulus which it has given to the development of the theory of numbers. It was his study of Fermat's last problem which led Kummer to the consideration of algebraic numbers and thus to the invention of his theory of ideals (compare p. 151). The problem itself and the various questions which have grown out of it engage at the present time the attention of many mathematicians all over the world; a large number of articles in the mathematical journals is devoted to them. These are quite beyond our reach. But we must try to get some notion of the simpler approaches to the question.

85. A new form of argument. In the first place we observe that if $(a, b; c)$ satisfies equation (8.1), then $(ka, kb; kc)$ will also be a solution of this equation; and conversely, if $(ka, kb; kc)$ is a solution, then $(a, b; c)$ is likewise a solution. These statements can be verified, just as the corresponding statement in Chapter I (compare p. 4), by direct substitution. Secondly, if two of the integers of a Fermat triple have a factor in common the third one must have

[1] Compare D. E. Smith, *A Source Book in Mathematics*, p. 213.

this factor as well. In virtue of these facts we can limit ourselves to the search for triples of integers which are relatively prime, two by two, and which satisfy Fermat's equation. We shall call them "*primitive Fermat triples.*" Finally, if n in equation (8.1) is the product of two numbers p and q, $n = pq$, and if $(a, b; c)$ is a solution of this equation, then $(a^p, b^p; c^p)$ is a solution of the equation $x^q + y^q = z^q$. A repetition of this argument leads to the conclusion that if Fermat's problem has *no* solution for any prime exponent greater than 2, then it can not have a solution for *any* exponent, unless all the factors of this exponent were equal to 2, i.e. unless $n = 2^k$. For instance, if we know that there exists no solution of the equation $x^7 + y^7 = z^7$, then we can conclude that there is no solution of the equation $x^{14} + y^{14} = z^{14}$. For if $(a, b; c)$ satisfied the latter equation then $(a^2, b^2; c^2)$ would satisfy the former. In general if $(a, b; c)$ were a solution of $x^{7k} + y^{7k} = z^{7k}$, then $(a^k, b^k; c^k)$ would be a solution of $x^7 + y^7 = z^7$.

There remain for consideration those equations (8.1) in which n is a prime number *and* those in which $n = 2^k$, $k > 1$. All numbers n of the latter kind contain the factor 4, so that we can put them in the form $4m$. The argument we have just made can be repeated to show that if $x^{4m} + y^{4m} = z^{4m}$ has a solution, then

$$(8.2) \qquad x^4 + y^4 = z^4$$

also has a solution; in other words, if (8.2) does not have a solution, then no equation of the form (8.1), in which $n = 4m$, can have a solution.

The study of equation (8.2) leads to some interesting facts which supplement in a useful way the results reached in Chapter I. We begin by establishing them.

Lemma 1. There exist no relatively prime integers u and v, of different parity, such that $u^2 + v^2$ and $u^2 - v^2$ are both equal to perfect squares.

Proof. If this lemma were not true, there would exist two relatively prime integers u and v of different parity such that $u^2 + v^2$ and $u^2 - v^2$ *are* perfect squares. Suppose then that

(8.3) $u^2 + v^2 = a^2$ and $u^2 - v^2 = b^2$, u and v relatively prime and of different parity. Then

$$(8.31) \qquad 2u^2 = a^2 + b^2 \quad \text{and}$$
$$(8.32) \qquad 2v^2 = a^2 - b^2,$$

while a and b are both odd. Consequently, since moreover $a > b$ as a result of (8.32), $a - b$ must be a positive *even* integer; therefore there must be a positive integer c such that $a - b = 2c$; in other words

$$a = b + 2c.$$

If we substitute this relation in (8.31) we find:

$$2u^2 = 2b^2 + 4bc + 4c^2, \quad \text{or} \quad u^2 = b^2 + 2bc + 2c^2 = (b + c)^2 + c^2.$$

But this means that $(b + c, c; u)$ is a Pythagorean triple. Moreover, if $b + c$ and c have a common factor k, i.e. if $b + c = kg$ and $c = kh$, then $a = b + 2c = k(g + h)$ and $b = b + c - c = k(g - h)$, so that we find, from (8.31) and (8.32), $2u^2 = 2k^2(g^2 + h^2)$ and $2v^2 = 4k^2gh$. This shows that u and v will then also have the factor k in common, which contradicts the hypothesis in (8.3). Therefore $b + c$ and c are relatively prime, so that $(b + c, c; u)$ is a *primitive* Pythagorean triple (compare p. 7). We can therefore conclude, in virtue of Theorem I, that there exist two more relatively prime integers of different parity, r and s, $r > s$ such that

$$(8.4) \qquad b + c = r^2 - s^2, \quad c = 2rs, \quad u = r^2 + s^2.$$

If we substitute those results in (8.32), we obtain

$$2v^2 = (a - b)(a + b) = 2c(2b + 2c) = 8rs(r^2 - s^2);$$

therefore $$v^2 = 4rs(r - s)(r + s).$$[1]

But since r and s are relatively prime and of different parity, no two of the four numbers r, s, $r - s$ and $r + s$ can have a factor in common (compare Lemma 3 on page 9, and remember that r and s are of different parity so that $r - s$ and $r + s$ are both odd). Their product can therefore be a perfect square only if each separately is a perfect square. Therefore there must exist four integers u_1, v_1, a_1 and b_1 such that $r = u_1^2$, $s = v_1^2$, $r + s = a_1^2$ and $r - s = b_1^2$; but this says that u_1, v_1, a_1, and b_1 satisfy the relations

$$(8.5) \qquad u_1^2 + v_1^2 = a_1^2 \quad \text{and} \quad u_1^2 - v_1^2 = b_1^2;$$

and, since r and s are relatively prime and of different parity, u_1 and v_1 must also be relatively prime and of different parity. We

[1] Instead of the first two of equations (8.4), we might have $c = r^2 - s^2$, $b + c = 2rs$. But since $2v^2 = 4c(b + c)$, we would be led to exactly the same expression for v^2 as was obtained above.

have come out of the underbrush into a place where we can look around a bit. For, equations (8.5) are of the same form as the equations (8.3) to which our lemma denies existence.

Apparently the discussion so far has not profited us very much. But the appearance is misleading! For, since $r > s$, we know that $2rs > 2s^2$, so that $a = b + 2c = r^2 - s^2 + 2rs > r^2 - s^2 + 2s^2 = r^2 + s^2$. If we remember now that it has been our tacit assumption throughout this argument that r, s, a and a_1 are all natural numbers, then we see at once $r^2 \geqq r$, $s^2 \geqq s$, and $a_1^2 \geqq a_1$. We are therefore justified in continuing the last statement as follows (notice that the properties of inequalities of page 28 again play an important part in the proceedings):

$$a > r^2 + s^2 \geqq r + s = a_1^2 \geqq a_1.$$

The state of affairs which we have now reached is therefore the following: If there is a set of four numbers u, v, a, b, which satisfies equations (8.3), then there is another set of four numbers u_1, v_1, a_1, b_1 which also satisfies these equations, and such that $a_1 < a$. But if this argument holds on the set u, v, a, b, it holds equally well for the set u_1, v_1, a_1, b_1. Therefore there would have to exist still another set u_2, v_2, a_2, b_2, which satisfies (8.3) and in which $a_2 < a_1$. On this set the argument could be repeated once again, so as to give a set u_3, v_3, a_3, b_3 with $a_3 < a_2$, and so forth. What would this lead to? Clearly to a denumerable infinitude of such sets, with a denumerable infinitude of steadily diminishing natural numbers a, a_1, a_2, a_3, ... But does an infinite set of diminishing natural numbers exist? Certainly not, for no diminishing collection of natural numbers can diminish lower than 1. It must therefore always have a last element; consequently it can not be infinite (compare Definition III, p. 13). Our argument based on the denial of Lemma 1 has thus led to a contradiction. In a logically organized world, such as we suppose our mathematical world to be, no contradictions can arise; therefore Lemma 1 must stand.

We have given this argument in full, partly because we wanted to establish the validity of the lemma, partly because it is simple although rather long, but chiefly because it is a nice example of a method of proof which is very frequently used in the theory of numbers. Its essential element is the demonstration that there would have to exist an infinitude of diminishing natural numbers if there existed a set of four integers which satisfy equations (8.3).

Whenever an argument can be made to lead to a descending infinitude of natural numbers the hypothesis upon which the argument rests becomes untenable. This method of proof is called the *method of infinite descent;* it was a favorite method with Fermat and it is still frequently employed. It would be interesting and valuable to compare this method with the method of mathematical induction (compare p. 91); but this does not fit in with present plans.

Lemma 1 clears the road for the excursion we want to make into the study of Fermat's last theorem. We deduce from it

Lemma 2. There exists no primitive Pythagorean triple $(p, q; r)$ in which p, q and r are all three perfect squares.

Proof. If such a triple did exist it would follow from Theorem I, that there are two relatively prime numbers u and v of different parity such that $u^2 + v^2 = r = a^2$, and $u^2 - v^2 = p$ (or q) $= b^2$. We have just proved that such numbers do not exist; therefore Lemma 2 has been established.

By means of Lemmas 1 and 2, we can now prove

Theorem XXVIII. There exists no solution in integers of the equation

$$x^4 + y^4 = z^4. \tag{8.2}$$

Proof. We already know that, if such a solution did exist, there would also be a primitive solution. In other words, there would exist a set of three integers, relatively prime two by two, for which the equation holds. If $(a, b; c)$ were such a triple, then $(a^2, b^2; c^2)$ would be a primitive Pythagorean triple. For to require that $(a^2, b^2; c^2)$ satisfy equation (1.1) and to require that $(a, b; c)$ satisfy (8.2) are equivalent questions; any answer to either furnishes an answer to the other. Lemma 2 gives us the assurance that there does not exist a primitive Pythagorean triple of the form $(a^2, b^2; c^2)$; consequently there is no solution of equation (8.2). This completes the proof of Theorem XXVIII.[1]

It will be a good thing to reread at this point the discussion of pages 167 and 168. For, in combination with Theorem XXVIII, it justifies the statement that the great Fermat theorem is equivalent to the assertion that no primitive Fermat triples exist for which n is a prime number different from 2. Indeed it is this assertion alone with which studies of the Fermat theorem are con-

[1] Compare also W. Lietzmann, *Der pythagoreische Lehrsatz,* pp. 65–68.

cerned. Our discussion has served to bring out what is essential in the statement of the theorem.

86. More impossibilities. The content of Lemma 1 can be put into a simple geometrical form different from the one suggested in Lemma 2. Direct geometrical interpretation of the equations (8.3) leads to the conclusion that if the legs and hypotenuse of two right triangles are measured by integers, then it can never happen that the two legs of one are equal to the hypotenuse and one leg of the other (see Fig. 27). Again, if the sides of two right triangles are related as those of triangles ABC and DBC in Fig. 27, then they can not all be measured by integers.

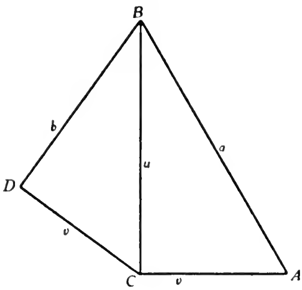


FIG. 27

Let us deduce now a few further consequences from our Lemma 1.

Theorem XXIX. There exists no Pythagorean triple such that the area of the corresponding right triangle is a perfect square.

Proof. If there were a non-primitive triple $(ka, kb; kc)$ for which the area of the corresponding triangle is a perfect square then the same would be true for the primitive triple $(a, b; c)$. For, since the area in the former case is $\frac{k^2ab}{2}$ and that in the latter case $\frac{ab}{2}$, it follows that either of those two numbers is a perfect square, provided the other is.

Suppose then that $(a, b; c)$ is a *p.P.t.* Theorem I assures the existence of two relatively prime numbers of different parity, r and s, $r > s$, such that

$$a = r^2 - s^2, \quad b = 2rs, \quad c = r^2 + s^2.$$

The area A of the corresponding triangle is then given by

$$A = \frac{ab}{2} = rs(r^2 - s^2).$$

If this were a perfect square, we would conclude, as in the proof of Lemma 1, that there is a set of four integers u_1, v_1, a_1, b_1 for which equations (8.5) hold (compare pp. 169 and 170); we have seen in Lemma 1 that this can not happen. Therefore A can not be a

perfect square. Notice that the proof of this theorem depends upon one of the essential elements in the proof of Lemma 1.

Theorem XXX. The equation $x^4 + y^4 = z^2$ can not be satisfied by three integers relatively prime in pairs.[1]

Proof. Suppose that $(a, b; c)$ were a solution of the equation

$$x^4 + y^4 = z^2, \tag{8.6}$$

and that a, b and c were relatively prime in pairs. Then $(a^2, b^2; c)$ would be a *p.P.t.* Hence there would exist two relatively prime integers u and v of different parity such that

$$a^2 = u^2 - v^2, \tag{8.61}$$

$$b^2 = 2uv, \tag{8.62}$$

$$c = u^2 + v^2. \tag{8.63}$$

If we write the first of these equations in the form $a^2 + v^2 = u^2$, we recognize that $(a, v; u)$ is a *p.P.t.*, so that *either* there are two relatively prime numbers r and s, of different parity, $r > s$ such that $a = 2rs$, $v = r^2 - s^2$, $u = r^2 + s^2$, *or* there are two such numbers for which

$$a = r^2 - s^2,\ v = 2rs,\ u = r^2 + s^2. \tag{8.7}$$

In the former case, we would conclude from (8.62) that $b^2 = 2(r^4 - s^4)$; but, since r and s are of different parity, r^4 and s^4 are also of different parity, so that $r^4 - s^4$ is odd. The right hand side of the last equation contains therefore only one factor 2; we conclude that it can not be a perfect square, and hence that this case can not arise.

In the second case, i.e. in case r, s, u, v and a satisfy the equations (8.7), we would derive from (8.62) the fact that $b^2 = 4rs(r^2 + s^2)$, and hence that $rs(r^2 + s^2)$ is a perfect square. Now r and s are relatively prime. If r and $r^2 + s^2$ had a common factor k, we could put $r = kr_1$ and $r^2 + s^2 = kt_1$. From this we would conclude that $s^2 = k(t_1 - kr_1^2)$, and hence that s^2 must have the factor k; since s^2 has no factors which s does not have, s would also have the factor k. But since r and s are relatively prime, this can not happen; ergo there are no factors common to any two of the three numbers r, s and $r^2 + s^2$. Hence their product can be a

[1] Compare Lietzmann, *op. cit.*, pp. 65–68.

perfect square only if each of them is a perfect square, i.e. if there exist three numbers, a_1, b_1 and c_1 relatively prime in pairs such that

$$r = a_1^2,\ s = b_1^2 \quad \text{and} \quad r^2 + s^2 = c_1^2;$$

that is to say, such that

$$a_1^4 + b_1^4 = c_1^2.$$

Thus we have arrived at a situation quite analogous to the one we encountered in the proof of Lemma 1 (compare p. 170), viz: From the existence of a solution $(a, b; c)$ of equation (8.6), we have deduced the existence of another solution $(a_1, b_1; c_1)$ of the same equation. Moreover we can show that $c_1 < c$. For from (8.63) and (8.7) we conclude that $c = u^2 + v^2 > u^2 \geqq u = r^2 + s^2 = c_1^2 > c_1$, by use of the same theorems on inequalities that were brought into play on page 170. Everything is now ready for the method of infinite descent to appear on the stage. For by repetition of the argument we have made so far we would obtain an infinitude of descending positive integers $c, c_1, c_2, \ldots$. There is then only one escape from contradiction, viz. the conclusion that no solution $(a, b; c)$ of equation (8.6) exists. This is the assertion of Theorem XXX.

Its content is clearly equivalent to the statement that no *p.P.t.* $(p, q; r)$ exists in which p and q are perfect squares.

Lemma 2 and hence Theorem XXVIII can be obtained by inference from Theorem XXX, instead of from Lemma 1. For, if there is no *p.P.t.* $(p, q; r)$ in which p and q are perfect squares, there can certainly not be one in which p, q and r are all three perfect squares. This second proof of Lemma 2 is not a waste of energy for us, because we are out not to see the landscape hastily but to become acquainted with its interlacing paths and with the methods of travel; not merely to hurry to the places frequented by tourists, but to get some insight into the connections which exist between those places. The new proof has given us another illustration of the power inherent in the method of infinite descent. Now let us see whether we can travel unaided a little farther in this part of the theory of numbers.

87. Some independent excursions.

1. Show that if any one of the equations $x^6 + y^6 = z^6$, $x^9 + y^9 = z^9$, $\ldots$ $x^{3k} + y^{3k} = z^{3k}$ possessed a solution in integers, then the equation $x^3 + y^3 = z^3$ would also have such a solution.

2. Determine all primitive Pythagorean triples for which the perimeter of the corresponding triangle is a perfect square (see Carmichael, *Diophantine Analysis*, p. 22, 3).

3. Show that if p and q are relatively prime and of different parity, $p > q$, then $pq(p^2 - q^2)(p^2 + q^2)$ can not be a perfect square.

4. Prove that the equation $x^4 - y^4 = z^2$ can not be satisfied by non-zero integers.

5. Prove that the equation $x^4 - 4y^4 = z^2$ can not be satisfied by non-zero integers.

6. Prove that the area of a primitive Pythagorean triangle can not be equal to twice a perfect square.

7. Prove that if $(p, q; r)$ is a *p.P.t.*, then neither p and r, nor q and r can be perfect squares.

8. Derive a proof of Lemma 2 from the result obtained in 7.

88. Apologia pro theoria numerorum. In the preceding sections we have seen something of the methods which are used in the Theory of Numbers, the part of mathematics which is concerned primarily with the properties of natural numbers. It has a long history; we have already mentioned the fact that Diophantus of Alexandria contributed to it about 1700 years ago.[1] During the past 20 years it has been developed in a very remarkable way. It has been extended into the theory of algebraic numbers, of which we have had a glimpse in Chapter VII, and intimate connections have been established between it and other fields of mathematics. Among those who have contributed a great deal to this development is G. H. Hardy, professor of mathematics at Cambridge University. At the conclusion of a lecture entitled "An introduction to the theory of numbers," delivered before the American Mathematical Society in December, 1928,[2] Professor Hardy said: "The elementary theory of numbers should be one of the very best subjects for early mathematical instruction. It demands very little previous knowledge; its subject matter is tangible and familiar; the processes of reasoning which it employs are simple, general and few; and it is unique among the mathematical sciences in its appeal to natural human curiosity. A month's intelligent instruction in the theory of numbers ought to be twice as instructive, twice as useful, and at least ten times as entertaining as the

[1] A monumental history of the subject is the work of a contemporary mathematician, L. E. Dickson, professor of mathematics at the University of Chicago, whose *History of the Theory of Numbers*, in 4 volumes, is the outstanding book in this field.

[2] See *Bulletin of The American Mathematical Society*, vol. 35, 1929, pp. 778–818.

same amount of 'calculus for engineers.' It is after all only a minority of us who are going to spend our lives in engineering workshops, and there is no particular reason why most of us should feel any overpowering interest in machines; nor is it in the least likely that, on those occasions when machines are of real importance to us, we shall require the power of dealing with them by methods more elaborate than the simplest rule of thumb. It is not engineering mathematics that is wanted for the understanding of modern physics and still less is it wanted by most of us for the ordinary needs of life; we do not actually drive cars by solving differential equations. There may be a case for subordinating mathematics to the linguistic and literary studies which are so much more obviously useful to ordinary men, but there is none for sacrificing a splendid subject to meet a quite imaginary need."

Because I am in hearty agreement with Professor Hardy's opinion, because it is a safe assumption that very few of my readers have had the advantage of "a month's intelligent instruction in the theory of numbers," and because it is hoped that the present chapter has aroused their "natural human curiosity," we are going to wander about for a short time in the garden of numbers. For those whose curiosity has been aroused to such an extent that a brief stroll can not satisfy them, there are more extensive guides available. Among them we mention the following: L. E. Dickson, *Introduction to the Theory of Numbers;* R. D. Carmichael, *The Theory of Numbers*. Both these books have been referred to earlier in this chapter. Another useful reference is the last named author's *Diophantine Analysis*.

89. New words for familiar ideas. In Chapter I we became acquainted with the concepts of "prime number" and of "pair of relatively prime numbers." A large part of the theory of numbers is concerned with these concepts. It would not be unnatural to think that, since as we go farther in the scale of natural numbers the number of available factors gets larger and larger, every number beyond a sufficiently large one would have at least one factor; in other words that the set of prime numbers forms a finite set. Nevertheless, the guess would be wrong. This was proved by Euclid[1] in the following very simple manner.

Theorem XXXI. The set of prime numbers forms an infinite set.

Proof. It follows from Definition III, that all we have to show

[1] See Heath, *A Manual of Greek Mathematics*, p. 240.

is that there is no *last* prime number. This Euclid did by the indirect method (compare p. 109). Suppose namely that 2, 3, 5, 7, 11, 13, . . . p represented the entire set of prime numbers. What would then be the factors of the number N obtained by multiplying all these prime numbers together and increasing the result by 1? This number N certainly exceeds p. Moreover it leaves a remainder 1 upon division by any of the numbers 2, 3, 5, 7, 11, 13, 17, . . . p; hence the number

$$N = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot \ldots \cdot p + 1$$

is divisible neither by 2, nor by 3, nor by 5, nor by 7, etc., and not by p. But if it were divisible at all, it would surely be divisible by a prime number. Hence either it is itself a prime number, or else it is divisible by a prime number *greater* than p. Neither of these alternatives is compatible with the assumption that p is the last prime number. Conclusion — there is no last prime number. Notice the simplicity of the argument, and the small amount of previous knowledge it requires (compare pp. 175, 176).

The prime numbers constitute then an infinite set. Because they are all contained among the natural numbers, this set is denumerable. It is therefore possible and not meaningless to ask for the 500th prime, or the 5000th prime. What those are I do not know. A person endowed with sufficient patience and industry, a long enough life and defective judgment as to the best way to use his time might find out what they are. When he had accomplished the task, we would ask him for the 5,000,000th prime; it is to be hoped that he would then see the point of the joke. What would be interesting and perhaps useful is a formula which gives the nth prime, i.e. a formula which would establish effectively the 1–1 correspondence which we know to exist between the set of prime numbers and the set of natural numbers. There are reasons to believe that no such formula can be expected; at any rate mathematicians are not interested in a search for it. For certain investigations in the theory of numbers it *is* important to have information available concerning large primes and concerning the least prime factors of large numbers. Such information has been collected in tables of primes and in factor tables. One of the most recent and best of such tables has been made by D. N. Lehmer,[1]

[1] See D. N. Lehmer, *Factor Tables for the First Ten Millions;* and *List of Prime Numbers from 1 to 10,006,721* by the same author.

who has devised very ingenious methods for detecting prime numbers and factors of large numbers. Very recently a machine was constructed by D. H. Lehmer[1] which makes it possible to reduce enormously the amount of time needed for such calculations.

Divisibility is a topic of primary importance in the theory of numbers. In the elementary school one learns how to tell whether a number is divisible by 2, 3, or 5; for divisibility by other prime numbers the tests become too complicated for practical use. Of most interest for us is the method by which they are established; to illustrate this method we shall consider a few simple cases.

Theorem XXXII. A number is divisible by 11 if and only if the sum of the digits in the even places and the sum of the digits in the odd places differ from each other by a multiple of 11.

Proof. (a) We will show first that every even power of 10 exceeds a multiple of 11 by 1; i.e. that 10^{2n} has the form $11p + 1$, or, that $10^{2n} - 1$ is always divisible by 11. This is obvious if it is observed that 10^{2n} consists of 1 followed by an even number of zeros, so that $10^{2n} - 1$ consists of an even number of nines and is therefore divisible by 11. For example $10^6 - 1 = 1{,}000{,}000 - 1 = 999{,}999 = 11 \times 90909$; hence $10^6 = 90909 \times 11 + 1$.

(b) Any odd power of 10 is equal to a multiple of 11 diminished by 1, i.e.

$$10^{2n+1} \text{ has the form } 11q - 1.$$

This fact *could* be proved by the same kind of reasoning as was used in part (a). It is a bit simpler and more satisfying to the æsthetic sense to establish it as a consequence of what has already been proved. This can be done readily; for

$$\begin{aligned} 10^{2n+1} = 10 \cdot 10^{2n} = 10 \cdot (11p + 1) &= 10p \cdot 11 + 10 \\ &= (10p + 1)\, 11 - 1. \end{aligned}$$

(c) It only remains to see that, if a number N is written in the decimal system, it is really represented as a sum of multiples of powers of 10. For instance, $72{,}059 = 7.10^4 + 2.10^3 + 5.10 + 9$. In language which is now familiar to us, we say that any number, written in the decimal scale, is represented as a polynomial in 10, with positive integral coefficients, *each of them less than* 10. Hence,

[1] See D. H. Lehmer, "A Photo-Electric Number Sieve," *American Mathematical Monthly*, vol. 40, 1933, p. 401.

the general representation of a number N in the decimal scale is as follows:

$$(8.8)\quad N = a_k 10^k + a_{k-1}10^{k-1} + \cdots + a_2 10^2 + a_1 10 + a_0,$$

where each of the coefficients $a_0, a_1, \ldots a_k$ is any one of the digits 0, 1, 2, . . ., 8 or 9. The digit a_0 stands in the first place, a_1 in the second, a_2 in the third; in general a_i is the digit in the $(i + 1)$th place. Hence $a_0, a_2, a_4 \ldots$ are the digits in the *odd* places; they are coefficients of *even* powers of 10. Similarly $a_1, a_3, a_5 \ldots$ are the digits in the even places; they are coefficients of *odd* powers of 10. The conclusion of the theorem should now be obvious: Since 10^{2n} has the form $11p + 1$, it follows that any term $a_{2n}10^{2n}$ is equal to a multiple of 11 *increased* by a_{2n}; and, similarly, any term $a_{2n+1}10^{2n+1}$ is equal to a multiple of 11 diminished by a_{2n+1}. Consequently N consists of a multiple of 11, increased by the sum of the digits in the odd places and decreased by the sum of the digits in the even places; it will therefore be divisible by 11 if and only if these two sums differ by a multiple of 11. This was the assertion of the theorem.

Take, for example, the number $N = 7{,}840{,}625$. The procedure of the theorem applied to this number leads to the following steps:

$$\begin{aligned} N &= 7\cdot 10^6 + 8\cdot 10^5 + 4\cdot 10^4 + 0\cdot 10^3 + 6\cdot 10^2 + 2\cdot 10 + 5 \\ &= 7(90909 \times 11 + 1) + 8(9091 \times 11 - 1) + 4(909 \times 11 + 1) \\ &\qquad + 6(9 \times 11 + 1) + 2(11 - 1) + 5 \\ &= [7 \times 90909 + 8 \times 9091 + 4 \times 909 + 6 \times 9 + 2]11 + 7 \\ &\qquad + 4 + 6 + 5 - (8 + 2). \end{aligned}$$

Hence 7,840,625 differs from a multiple of 11 by an amount equal to the difference between the sum of the digits in its odd places and the sum of the digits in its even places; this difference turns out to be 12. Therefore 7,840,625 is not divisible by 11.

The reader will have been impressed by a certain awkwardness in the discussion of Theorem XXXII. It arises because we used language not well adapted to the ideas we had to express. This inadequacy of our language would come out more sharply if we were to deal with something more complicated, as for instance divisibility by 7. To remedy the defect, we introduce some new concepts and words.

Definition XXXV. Two numbers a and b are called *congruent* with respect to a third number p if the difference between a and b is divisible by p.

The terminology actually used in the theory of numbers in place of "congruent with respect to p" is "congruent modulo p," and the notation is: $\equiv (\text{mod. } p)$; p is called the *modulus*.[1] Thus the formula $a \equiv b(\text{mod. } p)$ means that $a - b$ is divisible by p; it is called a *congruence.* For instance, $100 \equiv 1(\text{mod. } 11)$, and $30 \equiv 2(\text{mod. } 7)$ are congruences.

If $a \equiv b(\text{mod. } p)$, and $b \geqq 0$ but less than p, then b is called the residue of $a(\text{mod. } p)$.

Definition XXXVI. The *residue* of a, mod. p, is a non-negative number b, less than p, and such that $a \equiv b(\text{mod. } p)$.

For example the residue of 100, mod. 11 is 1, the residue of 100, mod. 7 is 2, etc. A few simple properties of congruences will be useful for our purpose. The discussion will familiarize us with the new concepts.

Theorem XXXIII. If $a \equiv b(\text{mod. } p)$ and $b \equiv c(\text{mod. } p)$ then $a \equiv c(\text{mod. } p)$.

Proof. The hypothesis means that $a - b = kp$, and $b - c = k_1p$; addition gives $a - c = (k + k_1)p$ which shows that, in accordance with the statement in the theorem, $a - c$ is divisible by p.

Theorem XXXIV. If $a \equiv b(\text{mod. } p)$, then $ac \equiv bc(\text{mod. } p)$. The proof can safely be left to the reader.

Theorem XXXV. If $a \equiv b(\text{mod. } p)$ and $a_1 \equiv b_1(\text{mod. } p)$, then $a + a_1 \equiv b + b_1(\text{mod. } p)$, and $aa_1 \equiv bb_1(\text{mod. } p)$.

Proof. In accordance with the hypothesis and Definition XXXV, we know that $a - b = kp$, and $a_1 - b_1 = k_1p$. Addition of these equations gives $a + a_1 - (b + b_1) = (k + k_1)p$, which establishes the first part of the theorem.

If we write the hypotheses in the form $a = b + kp$ and $a_1 = b_1 + k_1p$, and then multiply these equations, we find that $aa_1 = bb_1 + bk_1p + b_1kp + kk_1p^2 = bb_1 + (bk_1 + b_1k + kk_1p)p$, and hence that $aa_1 - bb_1 = (bk_1 + b_1k + kk_1p)p$; this establishes the second part of the theorem.

Using our new vocabulary, we can put the essential content of Theorem XXXII in the following statement: Any number is congruent, modulo 11, to the difference between the sum of the

[1] It is hardly necessary to observe that this use of the word "modulus" is entirely distinct from that which was brought forward in Chapter V.

digits in the odd places and the sum of the digits in the even places.

We turn now to divisibility by 7. The following congruences can be established readily without much calculation:

$$(8.9)\quad 10 \equiv 3(\text{mod. } 7),\quad 10^2 \equiv 2(\text{mod. } 7),\quad 10^3 \equiv 6(\text{mod. } 7),$$
$$10^4 \equiv 4(\text{mod. } 7),\quad 10^5 \equiv 5(\text{mod. } 7),\quad 10^6 \equiv 1(\text{mod. } 7).$$

The first of these is evident; from it we derive by means of Theorem XXXV that $10^2 \equiv 9(\text{mod. } 7)$, and hence, since $9 \equiv 2(\text{mod. } 7)$, by Theorem XXXIII that $10^2 \equiv 2(\text{mod. } 7)$. It will be easy for the reader to prove the remaining congruences.

Suppose now that we have a 7 digit number N, for instance 9,832,504. Then, in accordance with (8.8) we write

$$N = 9.10^6 + 8.10^5 + 3.10^4 + 2.10^3 + 5.10^2 + 4.$$

By means of the congruences (8.9), and Theorems XXXIV and XXXV, it follows then that

$$N \equiv 9 \times 1 + 8 \times 5 + 3 \times 4 + 2 \times 6 + 5 \times 2 + 4 = 87 \equiv 3(\text{mod. } 7),$$

which informs us that 9,832,504 will leave 3 as a remainder, upon division by 7.

The same treatment can be applied to any number of at most 7 digits. To deal with numbers of more than 7 digits, we derive from (8.9), by use of Theorem XXXV, the further congruences, mod. 7,

$$10^7 \equiv 3,\quad 10^8 \equiv 2,\quad 10^9 \equiv 6,\quad 10^{10} \equiv 4,\quad 10^{11} \equiv 5,\quad 10^{12} \equiv 1;$$
$$10^{13} \equiv 3,\quad 10^{14} \equiv 2,\quad 10^{15} \equiv 6,\quad 10^{16} \equiv 4,\quad 10^{17} \equiv 5,\quad 10^{18} \equiv 1;\ \text{etc.}$$

The fact that the residues, mod. 7, of the successive powers of 10 repeat themselves makes it possible to state a general test for divisibility by 7. Its essence is contained in the set of congruences (8.9); with those we shall be satisfied. It should be clear that a corresponding set of congruences for any other modulus, say 13, or 17, or any prime number p, lies at the basis of tests for divisibility by these numbers. The questions that must push themselves forward in the mind of the reader who has read this discussion, are these: If p is a prime number, will there always be a power of 10 whose residue, mod. p, is equal to 1? To how high a power must 10 be raised before the residue 1 is reached? A partial answer

to these questions is given in a later section — but it is time to do some exercises.

90. The use of an adequate language.

1. Formulate a test for divisibility by 7.
2. Develop the tests for divisibility by 3 and by 9.
3. Determine the residues, mod. 13, of the successive powers of 10.
4. Use the results of 3 to formulate a test for divisibility by 13.
5. Prove that, if $a \equiv b$, mod. n, and $n \equiv 0$, mod. p, then $a \equiv b$, mod. p.
6. Show that every number N is congruent, mod. p, to some number whose numerical value does not exceed $\frac{p}{2}$.
7. Prove that if $a \equiv b(\text{mod. } p)$, and $a \equiv -b(\text{mod. } p)$, then $a \equiv 0(\text{mod. } p)$, provided p is odd.
8. Prove that if $a^2 \equiv b^2(\text{mod. } p)$, and p is a prime, then at least one of the congruence $a \equiv b$ and $a \equiv -b$ must hold modulo p.
9. Determine the residues, mod. 7, of the powers of 9 and prove the formula Res. $9^n(\text{mod. } 7) \equiv 2^{\text{Res. } n(\text{mod. } 3)}$.
10. Prove (1) that if $a_1, a_2, a_3, \ldots a_p$ is any set of p consecutive integers, then the residues, mod. p of the numbers of this set constitute an arrangement of the numbers $0, 1, 2, \ldots p - 1$; (2) that in a set of p consecutive integers, there is always exactly one multiple of p.
11. Prove (1) that if $a_1, a_2, a_3, \ldots a_p$ is a set of p consecutive terms of an arithmetic progression whose difference is relatively prime to p, then the residues, mod. p, of the numbers of this set constitute an arrangement of the numbers $0, 1, 2, \ldots p - 1$; (2) that, among p consecutive terms of an arithmetic progression whose difference is relatively prime to p, there is always exactly one multiple of p.
12. Determine which of the field properties hold for the set of residues, mod. p, of the natural numbers, if all sums and products are replaced by their residues, mod. p (compare *25*, *26*).

91. Something from the theory of numbers. Before taking up the questions raised at the end of *89*, we want to have a look at one of the early attempts to obtain a formula for prime numbers. It is contained in a theorem known as Wilson's Theorem [1]; its enunciation is helped by the use of the following technical term:

Definition XXXVII. The product of the integers 1, 2, 3, 4, . . . n is called *n factorial* and is denoted by $n!$

[1] The theorem is named after the Englishman John Wilson (1741–1793); it seems to have been proved for the first time in 1771, by the famous French mathematician Joseph Louis Lagrange (1736–1813).

For instance $5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$; $7! = 5! \cdot 6 \cdot 7 = 5040$. We can now state Wilson's theorem in the following way:

Theorem XXXVI. If and only if p is a prime number, then $(p-1)! + 1 \equiv 0 (\text{mod. } p)$.[1]

Proof. (a) Let us consider the integers

$$(8.10) \qquad 1, 2, 3, \ldots p-1,$$

the number p being a prime. Since they are all less than p, none of them is divisible by p. Along with them we consider the integers

$$(8.101) \qquad 2, 4, 6, \ldots 2p-4, 2p-2,$$

obtained from those in the set (8.10) by multiplying them by 2, supposing $p > 2$. The residues, mod. p, of the numbers in (8.101) must all appear in (8.10). Moreover no two numbers in (8.101) can have the same residue, mod. p. For if $b_1 = 2a_1$ and $b_2 = 2a_2$, $b_1 > b_2$, are two numbers in (8.101) with the same residue, mod. p, then $b_1 - b_2 = 2(a_1 - a_2) \equiv 0(\text{mod. } p)$, where a_1 and a_2 belong in (8.10), $a_1 > a_2$. This means that $2(a_1 - a_2)$ is divisible by p. But $a_1 - a_2$, being a number of the set (8.10), is not divisible by p; since moreover $p > 2$, $2(a_1 - a_2)$ can not be divisible by p. Therefore no two numbers in (8.101) can have the same residue, mod. p. Consequently, the set of residues of the numbers in (8.101) is the same as the set of numbers (8.10), but rearranged in some way.[2] There is therefore, for each number a in (8.10), exactly one number b in (8.101) whose residue is a. In particular there is one number in (8.101) whose residue is equal to 1; i.e. there is exactly one number in (8.10), let us call it a_1, such that $2a_1 \equiv 1(\text{mod. } p)$; and clearly, $a_1 \neq 1$.

By an exactly analogous reasoning, we show that if $p > 3$ the set $3, 6, 9, \ldots 3p-6, 3p-3$ contains exactly one number which has the residue 1, mod. p; i.e. that there exists exactly one number in (8.10), let us call it a_2, such that $3a_2 \equiv 1(\text{mod. } p)$; we have again $a_2 \neq 1$. In this way we can continue, until we have obtained $p-2$ numbers $a_1, a_2, a_3, \ldots a_{p-2}$ from the set (8.10) none of them equal to 1 and such that

$$(8.102) \qquad 2a_1 \equiv 3a_2 \equiv 4a_3 \equiv \cdots \equiv (p-2)a_{p-3} \equiv (p-1)a_{p-2} \equiv 1(\text{mod. } p).$$

[1] Illustration: For $p = 7$, we find indeed that $6! + 1 = 721 \equiv 0(\text{mod. } 7)$; but for $p = 8$, we have $7! + 1 = 5041 \equiv 1(\text{mod. } 8)$.

[2] Compare *90*, 11.

But no two of these numbers can be the same, for, if a_i were equal to a_j, we would have besides the relation $(i + 1)a_i \equiv 1(\text{mod. } p)$, also the fact that $(j + 1)a_i \equiv 1$, and hence that $(i - j)a_i \equiv 0(\text{mod. } p)$. The argument used in the discussion of (8.101) shows that this is impossible.

The set of numbers $a_1, a_2, \ldots a_{p-2}$ constitutes therefore a rearrangement of the numbers $2, 3, \ldots p - 1$, and the products in (8.102) are products of pairs of numbers selected from the set (8.10).

(b) Is there any pair among them whose two factors are the same? If $a \cdot a$ were such a pair, we would have $a^2 \equiv 1(\text{mod. } p)$, or $(a - 1)(a + 1) \equiv 0(\text{mod. } p)$. Now, since a is at most $p - 1$, and since p is a prime, $a - 1$ can not have any factor in common with p, and if $a - 1 \neq 0$, it follows that $a + 1$ can only be p. The congruence $a^2 \equiv 1(\text{mod. } p)$ leads therefore to the two possibilities $a - 1 = 0$ or $a + 1 = p$, i.e. $a = 1$, or $a = p - 1$. But the set $a_1, a_2, \ldots a_{p-2}$ does not contain 1. Therefore the only product in (8.102) which can be a square is $(p - 1)a_{p-2}$ i.e. $(p - 1)^2$; and indeed $(p - 1)^2 = p^2 - 2p + 1 \equiv 1(\text{mod. } p)$. There remain then the $p - 3$ products $2a_1, 3a_2, 4a_3, \ldots (p - 2)a_{p-3}$ each of which consists of two distinct factors and each of which is congruent to 1, mod. p; moreover, the set $a_1, a_2, \ldots a_{p-3}$ is a rearrangement of the set $2, 3, \ldots p - 2$. We can therefore select one half of this set of products[1] in such a way that each number of the set $2, 3, \ldots p - 2$ occurs once and only once as a factor in one of these pairs. Since each of the products in this half of the set has the residue 1, mod. p., their product is, in virtue of Theorem XXV, also congruent to 1, mod. p. Therefore

$$2 \cdot 3 \cdot 4 \cdot \ldots \cdot (p - 3)(p - 2) \equiv 1(\text{mod. } p).$$

From this we derive, by means of Theorem XXXIV (remember also Definition XXXVII) the consequence that

$$(p - 1)! \equiv p - 1(\text{mod. } p),$$

and hence that $(p - 1)! + 1 \equiv p \equiv 0(\text{mod. } p)$.

This completes the proof of the first part of Wilson's theorem.

(c) The second part states that if p is a composite number, then $(p - 1)! + 1$ is not congruent to 0, mod. p. It is immediately evident that this is so, because if p has a factor p_1, then $(p - 1)!$

[1] Notice that, since p is a prime > 2, $p - 3$ is an even number.

is divisible by p_1. Therefore $(p - 1)! + 1$ leaves a remainder 1 upon division by p_1. Hence it is not divisible by p_1 and certainly not by p.

It will be instructive to follow the steps of the proof in a particular case. Let us do this for $p = 13$. The set (8.10) then becomes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12; and the set (8.101) is 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24. The residues mod. 13 of these numbers are, in order, equal to 2, 4, 6, 8, 10, 12, 1, 3, 5, 7, 9, 11, which is indeed a rearrangement of the numbers from 1 to 12; the one number in the second set whose residue equals 1, is 14; hence $a_1 = 7$.

The residues mod. 13 of the numbers 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36 are equal to 3, 6, 9, 12, 2, 5, 8, 11, 1, 4, 7, 10. Since $27 \equiv 1 (\text{mod. } 13)$, we find that $a_2 = 9$.

Continuing in this way, we find $a_3 = 10$, $a_4 = 8$, $a_5 = 11$, $a_6 = 2$, $a_7 = 5$, $a_8 = 3$, $a_9 = 4$, $a_{10} = 6$, $a_{11} = 12$. The products (8.102) are therefore

$$2 \cdot 7, 3 \cdot 9, 4 \cdot 10, 5 \cdot 8, 6 \cdot 11, 7 \cdot 2, 8 \cdot 5, 9 \cdot 3, 10 \cdot 4, 11 \cdot 6, 12 \cdot 12.$$

All of these products have the residue 1, mod. 13, and only the last one is a perfect square. The product of the first 5 is

$$2 \cdot 7 \cdot 3 \cdot 9 \cdot 4 \cdot 10 \cdot 5 \cdot 8 \cdot 6 \cdot 11 = 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 \cdot 11.$$

It must be congruent to 1, mod. 13; therefore $12! \equiv 12 (\text{mod. } 13)$ and $12! + 1 \equiv 13 \equiv 0 (\text{mod. } 13)$.

Theoretically Wilson's theorem provides a perfect test for prime numbers. To determine whether a given number N is a prime, we have but to calculate $(N - 1)! + 1$ and to examine whether or not the result is divisible by N. But this calculation and the following examination would be a stupendous task as soon as N gets larger than 20 or 25. This test is therefore not of much value for the actual determination of large prime numbers.

As a by-product of the proof of this theorem, we obtain an answer to the questions asked at the end of *89*. For, we have seen, that if p is a prime, and $k < p$, then the residues, mod. p, of the set of numbers

$$(8.11) \qquad k, 2k, 3k, \ldots (p - 1)k,$$

taken in order, constitute a rearrangement of the numbers 1, 2, 3, $\ldots p - 1$. It is easy to see that this conclusion remains valid

even if $k > p$, provided only that k is not divisible by p. For the residues, mod. p, of the numbers in (8.11) must even then always be contained among the numbers 1, 2, 3, . . . $p - 1$, since none of them is divisible by p. Moreover no two of them can have the same residue, because in that case they would be congruent, mod. p. And if $ak \equiv bk$, mod. p, while $a < p$, $b < p$, then $(a - b)k$ would be divisible by p, which is impossible, if p is a prime and k is not divisible by p. But if each of the numbers in (8.11) is congruent to one number in (8.10) and vice versa then the product of the numbers in the first set is, by Theorem XXXV, congruent to the product of the numbers in the second set; that is to say:

$$(p - 1)!\, k^{p-1} \equiv (p - 1)!\ (\text{mod. } p),$$

or

$$(p - 1)!\,(k^{p-1} - 1) \equiv 0 (\text{mod. } p).$$

But $(p - 1)!$ has no factor in common with p, since p is a prime. Therefore $k^{p-1} - 1$ must be divisible by p; hence

$$k^{p-1} - 1 \equiv 0 (\text{mod. } p).$$

This interesting formula was discovered by Fermat (compare p. 166) and is usually referred to as Fermat's theorem.

Theorem XXXVII. If p is a prime and k is any number which is not divisible by p, then $k^{p-1} - 1$ is divisible by p.

For example, taking $p = 5$ and $k = 4$, we find that $4^4 - 1 = 256 - 1$ is indeed divisible by 5. It is clear that if k is divisible by p, then $k^N \equiv 0 (\text{mod. } p)$ no matter what integer N is.

The answer to the question: "If p is a prime, will there always be a power of 10 whose residue, mod. p, is equal to 1?" is therefore: "*No*, if $p = 2$ or 5, and *Yes*, if p is any other prime." As to the question: "To how high a power must 10 be raised before the residue 1 is reached," this can not be definitely answered. We do know that if $p \neq 2$ and 5, then $10^{p-1} \equiv 1 (\text{mod. } p)$; e.g. $10^6 \equiv 1$, mod. 7; and $10^{16} \equiv 1$ mod. 17. But we do not know whether the residue 1 can never be obtained when 10 is raised to a power whose exponent is less than $p - 1$.[1]

92. To antiquity once more. The reader has now had a few samples from the elementary theory of numbers. They do not give him the "month's intelligent instruction" recommended by G. H. Hardy, but they should give him some idea of the sort of thing such instruction would be occupied with. He would learn

[1] Compare Dickson, L. E., *Introduction to the Theory of Numbers*, Chap. II.

a good deal more about congruences of numbers; he would learn how to determine numbers x such that $ax + b \equiv 0 (\text{mod. } p)$, he would hear about quadratic residues and about many things that the usual undergraduate college course does not take up, but that are easily within reach of any one gifted with normal intelligence. All these things we have to leave untouched; in accordance with our general plan we do not go in for the more technical parts of the subject. But it would be a mistake to leave this topic without the consideration of a simple but important contribution made in antiquity, which has become the cornerstone of a vast structure that spreads not only over the theory of numbers but over several other parts of mathematics.

The monument which we are going to visit is the method devised by Euclid for finding the *greatest common divisor* (g.c.d.) of two numbers. Euclid is known, at least by name, to every schoolboy who has studied plane geometry. It is to him that the systematic exposition of the geometrical knowledge acquired by Greek mathematicians is usually attributed.[1] His famous *Elements*, and several of his other writings, contain a good many results which are not usually classified as topics in geometry. Among them is his proof that the set of prime numbers is an infinite set (compare p. 176). Another one is his method for finding the g.c.d. of two numbers.

Lemma 1. If, in the set of natural numbers, a and b both possess the factor k, then, for every pair of natural numbers p and q, the numbers $pa + qb$ and $pa - qb$ also have the factor k; it is understood, in order to remain within the set of natural numbers, that we must have $pa > qb$.

Proof. If $a = ka_1$, and $b = kb_1$, then $pa + qb = pka_1 + qkb_1 = k(pa_1 + qb_1)$; and $pa - qb = k(pa_1 - qb_1)$; thus the Lemma is proved. The reader will do well to ascertain which of the properties of the set of natural numbers are here used. (Compare 25).

Corollary. If, two natural numbers a and b being given, there exist two natural numbers p and q such that $pa - qb$, or $qb - pa$, is equal to 1, then a and b are relatively prime.

The proof of this corollary can safely be left to the reader.

Lemma 2. If a and b are natural numbers, $a > b$, then there exist two other natural numbers, q and r, $0 \leqq r < b$, such that $a = qb + r$.

[1] Compare, e.g., Heath, *A Manual of Greek Mathematics*, pp. 204–255, in particular p. 236; also Cajori, *A History of Mathematics*, pp. 35–40.

Proof. Consider the multiples of b, viz. b, $2b$, $3b$, $4b$, etc. . . . Since $a > b$, there must be among them two successive multiples of b, such that a either lies between them or else coincides with the smaller of the two. (Why must two such multiples of b exist? Compare the footnote on page 260, concerning the Axiom of Archimedes.)

Suppose they are qb and $(q + 1)b$; then

$$qb \leqq a < (q + 1)b.$$

Hence, *either* $a = qb$, *or* $a = qb + r$, $0 < r < b$. In either case we have $a = qb + r$, $0 \leqq r < b$. The process by which the numbers q and r are determined, is called division, q is called the quotient, r the remainder. When for example we find that 517, upon division by 23, gives a quotient of 22 and leaves a remainder of 11, we have found that $517 = 22 \cdot 23 + 11$.

In the terminology of the present chapter, we would say that $r =$ residue a(mod. b). We have repeatedly (and tacitly) assumed in the preceding sections that the determination of residues is always possible. The significance of Lemma 2 lies in its attempt to prove this assumption!

Lemma 3. The g.c.d. of two numbers a and b, $a > b$, is equal to the g.c.d. of b and r, $b > r \geqq 0$, where r is the residue of a, mod. b.

Proof. We know from Lemma 2, that numbers q and r exist, such that $0 \leqq r < b$ and $a = qb + r$; hence, it follows from Lemma 1, that any divisor of b and r is also a divisor of a and therefore of a and b. But we can also write $r = a - qb$, so that we now conclude, by use of Lemma 1, that any factor common to a and b is also a factor of r and therefore a common divisor of b and r. Since any common divisor of a and b is a common divisor of b and r; and any common divisor of b and r is also a common divisor of a and b, the g.c.d. of a and b is equal to the g.c.d. of b and r. In case $r = 0$, the g.c.d. of b and r and hence the g.c.d. of a and b is b itself; this is evident a priori.

Lemma 4. If a and b are natural numbers, $a > b$, there exists a finite diminishing sequence of numbers $r, r_1, r_2, \ldots r_k$, such that r is the residue of a, mod. b; r_1 the residue of b, mod. r; r_2 the residue of r, mod. r_1, and so forth until we reach r_k as the remainder upon division of r_{k-2} by r_{k-1}; and such that r_k is an exact divisor of r_{k-1}.

Proof. In Lemma 2 we have already shown that, when $a > b$, there exists a number r, the remainder in the division of a by b,

such that $r < b$. We can now apply Lemma 2 to the numbers b and r, and obtain r_1 such that $r_1 < r$; applying it to r and r_1, we obtain r_2, such that $r_2 < r_1$; and so forth. Thus we obtain a diminishing sequence of natural numbers r, r_1, r_2, etc.; this sequence must be finite (compare p. 171). It stops as soon as a number r_k is reached which is an exact divisor of its precedessor r_{k-1}. Unless this happens at an earlier point, it will surely take place when we reach the number 1; in this case the final number r_k of the sequence is 1. The process can be stopped before this point is reached only if one of the divisions leaves no remainder. If the divisor in this division is d, then $r_k = d$ is the last element of the sequence.

Theorem XXXVIII. The g.c.d. of two numbers a and b, $a > b$, is the final element r_k in the sequence of numbers, whose existence was established in Lemma 4.

Proof. It follows from Lemma 3 and from the way in which the sequence of Lemma 4 was built up that

g.c.d. of a and b = g.c.d. of b and r = g.c.d. of r and r_1 = g.c.d. of r_1

and r_2 = . . . = g.c.d. of r_{k-1} and r_k. But, since r_k is an exact divisor of r_{k-1}, the g.c.d. of r_{k-1} and r_k is r_k. This completes the proof of our theorem.

The scheme which serves to determine r_k is called the *Euclidean algorithm* (compare p. 54). If $q, q_1, q_2, \ldots q_k$ are the quotients in the successive divisions we can write the Euclidean algorithm in the following form

$$a = qb + r,\ b = q_1 r + r_1,\ r = q_2 r_1 + r_2,\ r_1 = q_3 r_2 + r_3,\ \ldots$$

$$r_{k-3} = q_{k-1} r_{k-2} + r_{k-1},\ r_{k-2} = q_k r_{k-1} + r_k,\ r_{k-1} = q_{k+1} r_k. \tag{8.12}$$

Let us consider as an example the problem of finding the g.c.d. of 782 and 3315. By successive divisions, we find that the relations (8.12) are in this case as follows:

$$3315 = 4 \times 782 + 187;\ 782 = 4 \times 187 + 34;$$
$$187 = 5 \times 34 + 17;\ 34 = 2 \times 17.$$

In this case therefore $q = 4$, $r = 187$; $q_1 = 4$, $r_1 = 34$; $q_2 = 5$, $r_2 = 17$; $k = 2$; the g.c.d. of 782 and 3315 is 17. Indeed, we see that $782 = 17 \times 46$, and $3315 = 17 \times 195$, and that 46 and 195 have no common factors.

An interesting deduction can be made from the relations (8.12). If they be written, excepting the final relation $r_{k-1} = q_{k+1}r_k$, as sentences of which $r, r_1, r_2, \ldots r_k$ are the subjects, they take the form

$$r = a - qb,\quad r_1 = b - q_1 r,\quad r_2 = r - q_2 r_1,\quad r_3 = r_1 - q_3 r_2,\ \ldots$$
$$r_{k-1} = r_{k-3} - q_{k-1}r_{k-2},\ r_k = r_{k-2} - q_k r_{k-1}.$$

The first of these expresses r in terms of a and b. If, in the second, we replace r by the expression given for it in the first, we find $r_1 = b - q_1(a - qb) = (1 + qq_1)b - q_1 a$, which expresses r_1 in terms of a and b. We can now use this relation together with the first, in the third; this gives us $r_2 = a - qb - q_2[(1 + qq_1)b - q_1 a] = (1 + q_1 q_2)a - (q + q_2 + qq_1q_2)b$, which gives r_2 in terms of a and b.

We can write the results obtained so far in the following form:

$$\begin{array}{llll} r = A_0 a - B_0 b, & \text{when } A_0 = 1, & B_0 = q; \\ r_1 = B_1 b - A_1 a, & \text{“}\quad A_1 = q_1, & B_1 = 1 + qq_1; \\ r_2 = A_2 a - B_2 b, & \text{“}\quad A_2 = 1 + q_1 q_2, & B_2 = q + q_2 + qq_1q_2. \end{array}$$

If this process has been carried far enough to express $r, r_1, r_2, \ldots r_{i-2}, r_{i-1}$ in terms of a and b, then it can obviously be carried a step further. For we would then have (let us suppose that i is an even number — the argument is the same if i is odd, although the formulas look slightly differently):

$$r_{i-2} = A_{i-2}a - B_{i-2}b,\quad r_{i-1} = B_{i-1}b - A_{i-1}a;$$

and also
$$r_i = r_{i-2} - q_i r_{i-1}.$$

Substitution of the first two in the last would give

$$\begin{aligned} r_i &= A_{i-2}a - B_{i-2}b - q_i(B_{i-1}b - A_{i-1}a) \\ &= (A_{i-2} + q_i A_{i-1})a - (B_{i-2} + q_i B_{i-1})b = A_i a - B_i b, \end{aligned}$$

when we put

$$A_i = A_{i-2} + q_i A_{i-1},\quad B_i = B_{i-2} + q_i B_{i-1}.$$

The principle of mathematical induction [1] now justifies the conclusion that every one of the numbers $r, r_1, r_2, \ldots r_k$ can be expressed in terms of a and b.

We record the result as follows:

Theorem XXXIX. The g.c.d. of two numbers a and b can be expressed in the form $Aa - Bb$ or in the form $Bb - Aa$, where A and B are natural numbers.

[1] Compare pp. 91, 163.

In the numerical example which we considered it was found that $k = 2$; the g.c.d. is then $r_2 = A_2a - B_2b$, where $A_2 = 1 + q_1q_2$, $B_2 = q + q_2 + qq_1q_2$. If we use the values of q, q_1 and q_2 which were recorded above, we find

$$A_2 = 1 + 4 \times 5 = 21, \; B_2 = 4 + 5 + 4 \times 4 \times 5 = 89,$$

so that
$$r_2 = 21 \times 3315 - 89 \times 782;$$

it is easily verified that this leads indeed to the value 17 for the g.c.d. of 3315 and 782.

If the two numbers a and b are relatively prime, $r_k = 1$. We have then an interesting special case of our theorem.

Corollary. If a and b are two relatively prime numbers, there exist two natural numbers A and B such that either $Aa - Bb$ or else $Bb - Aa$ is equal to 1.

If $a = 15$ and $b = 8$, the relations (8.12) become $15 = 8 + 7$ and $8 = 7 + 1$; hence $7 = 15 - 8$, and $1 = 8 - 7 = 8 - (15 - 8) = 2 \times 8 - 1 \times 15$; so that $B = 2$, $A = 1$, and the second alternative applies. This last corollary has interesting applications in many parts of mathematics. Some indication of them can be found in the exercises which follow. But we can not forego the pleasure of discovering a path which connects the region we occupy now with one of the important scenes in the preceding chapter.

We notice in the first place that the corollary which has just been proved is the converse of the corollary which follows Lemma 1 on page 187. By combining them and using positive or negative integers instead of natural numbers for p and q, we obtain

Theorem XL. The natural numbers a and b are relatively prime *if and only if* there exist integers p and q such that $pa - qb = 1$.

Next we prove that if b is relatively prime to a_1 and to a_2, then it is also relatively prime to the product a_1a_2. For the hypothesis enables us to say, on the strength of Theorem XL, that there exist integers p_1 and q_1, such that $p_1a_1 - q_1b = 1$, and also integers p_2 and q_2 such that $p_2a_2 - q_2b = 1$. From these relations, written in the form $p_1a_1 = 1 + q_1b$ and $p_2a_2 = 1 + q_2b$, follows that $p_1p_2a_1a_2 = 1 + (q_1 + q_2 + q_1q_2b)b$. This shows that there exist integers p and q, viz. $p = p_1p_2$ and $q = q_1 + q_2 + q_1q_2b$ such that $pa_1a_2 - qb = 1$, and hence, by using Theorem XL once more, that b is relatively prime to a_1a_2.

Repetition of this argument leads to the conclusion that if the

natural number b is relatively prime to each of 3, or 4, or 5, . . ., or n natural numbers $a_1, a_2, a_3, \ldots a_n$, then it is also relatively prime to their product $a_1a_2a_3 \ldots a_n$.

Suppose now that the natural number N can be factored into prime factors in two ways, i.e. suppose there are two sets of *prime* numbers, $a_1, a_2, a_3, \ldots a_n$ and $b_1, b_2, b_3, \ldots b_m$ (some of the a's may be equal to each other, also some of the b's), such that

$$N = a_1a_2a_3 \ldots a_n = b_1b_2b_3 \ldots b_m. \tag{8.13}$$

Then b_1 is obviously a factor of the product $a_1a_2a_3 \ldots a_n$. But, if b_1 were relatively prime to each of the numbers $a_1, a_2, a_3, \ldots a_n$, then it would also be relatively prime to their product, as we have just seen. Therefore b_1 must have a factor in common with at least one of the a's. But b_1, as well as all the a's, is a prime number; hence b_1 must be equal to one of the a's. Let us suppose, for convenience, that $b_1 = a_1$. Then we obtain from (8.13) the equality

$$a_2a_3a_4 \ldots a_n = b_2b_3b_4 \ldots b_m.$$

By treating it as we have treated (8.13), we conclude that b_2 must be equal to one of the remaining a's. If we suppose $b_2 = a_2$, we are led to the equality $a_3a_4a_5 \ldots a_n = b_3b_4b_5 \ldots b_m$ with which we can proceed in the same manner. Continuing in this way, we reach as a final conclusion

Theorem XLI. A natural number can be factored into prime factors in one and only one way.

Thus we have obtained the conclusion which was stated, but not proved, in *74*. But, more important, we have established a connection between the Euclidean greatest common divisor algorithm and unique factorization. With the opening up of this road, a much traveled one in contemporary work, we close our first expedition in the foothills of the theory of numbers.

93. Excursions among integers.

1. Determine the residues, mod. 11, of the numbers obtained by multiplying the set of numbers from 1 to 10 by 2, by 3, by 4, and so forth up to and including 10.

2. Determine the numbers in each of the preceding sets which are congruent to 1, mod. 11. Use the results to verify the fact that $10! + 1 \equiv 0$, mod. 11.

3. Use the results of 1 to verify the fact that $5^{10} - 1 \equiv 0$, mod. 11.

4. Verify the fact that $15^{10} - 1 \equiv 0$, mod. 11, and that $15^{12} - 1 \equiv 0$, mod. 13.

5. Determine the g.c.d. of 1329 and 5724; express the result as the difference of multiples of the two given numbers.

6. Show that 2847 and 3592 are relatively prime numbers, and determine multiples of these numbers whose difference is equal to 1.

7. Solve the equation $5x + 7y = 1$ in integers.

8. Show that if $x = a$, $y = b$ is a solution in integers of the equation in the preceding problem, then $x = a + 7t$, $y = b - 5t$ is also such a solution, for every integral value of t.

9. Prove that if $x = a$, $y = b$ and $x = a_1$, $y = b_1$ are two solutions in integers of the equation $5x + 7y = 1$, then $a_1 - a : b_1 - b = 7 : -5$; in other words, that there exists then a number t such that $a_1 = a + 7t$, $b_1 = b - 5t$; or again that all the admissible values of x are congruent mod. 7, and all the admissible values of y congruent mod. 5.

10. Prove that if $x = a$, $y = b$ is any solution in integers of the equation $5x + 7y = 1$, then $x = a + 7t$, $y = b - 5t$ represents all its solutions in integers, if t is given all integral values.

Remark. The solution $x = a + 7t$, $y = b - 5t$ is then called the general solution in integers of the given equation.

11. Determine the general solution in integers of the equation $17x + 13y = 1$.

12. Deduce Theorem XXXIX from the Corollary on page 191.

13. From the result in 7, obtain first one solution, and then the general solution in integers of the equation $5x + 7y = 2$.

14. Obtain the general solution in integers of the equation $17x + 13y = 5$.

15. Prove that the equation $px + qy = r$, in which p, q and r are given integers, positive or negative, has no solution in integers if p and q have a common factor, *unless* r also has this factor.

16. Show that the equation $px + qy = r$ is of interest only if p and q are relatively prime.

17. Develop a general method for determining the general solution in integers of the equation $px + qy = r$, in which p and q are relatively prime. Apply this general method to the equations $4x - 7y = 18$, $6x + 25y = 7$, and $17x - 13y = 9$.

18. The sum of $100 is to be spent for books, some to cost $2 and others to cost $3. How many books of each kind can be bought?

19. Determine all positive integers which leave a remainder 3 upon division by 5, and a remainder 4 upon division by 7. What is the meaning of the word "all" in this problem?

20. Determine all numbers x which satisfy the following two congruences $x \equiv 5 (\text{mod. } 7)$ and $x \equiv 8 (\text{mod. } 11)$.

21. Prove that, if p and q are relatively prime, then $a \equiv b(\text{mod. } p)$, and $a \equiv b(\text{mod. } q)$ implies that $a \equiv b(\text{mod. } pq)$.

22. Prove that, if $a \equiv b(\text{mod. } pq)$, then $a \equiv b(\text{mod. } p)$ and $a \equiv b(\text{mod. } q)$.

23. Determine the g.c.d. of 7,548 and 14,245.

24. Which positive numbers are congruent to 1, mod. 2, mod. 5 and mod. 7?

CHAPTER IX

SOME AMUSEMENT AND SOMETHING ELSE

The notary then took it upon himself to justify Mademoiselle Préfère's educational system, and observed by way of conclusion, "It is not by amusing oneself that one learns." "It is only by amusing oneself that one can learn," I replied. — Anatole France, *The Crime of Sylvester Bonnard* (translation by Lafcadio Hearn).

94. Major and minor scales. The concepts with which we have become acquainted in the preceding chapter give rise to a number of developments, of interest in themselves and some of them of importance in different fields of mathematics. Let us begin by considering one of the latter kind.

In *89* (see p. 179) it was observed that, in the familiar decimal scale of notation, the number $N = 7{,}840{,}625$ represents $7 \cdot 10^6 + 8 \cdot 10^5 + 4 \cdot 10^4 + 0 \cdot 10^3 + 6 \cdot 10^2 + 2 \cdot 10 + 5$, and that, in general, a number whose digits are $a_k, \ldots a_3, a_2, a_1, a_0$ stands for $a_k 10^k + \ldots + a_3 10^3 + a_2 10^2 + a_1 10 + a_0$. Conversely, every polynomial in 10, whose coefficients are selected from the digits 0, 1, 2, . . . 9 is represented in the decimal scale by a number whose digits are the coefficients.

The name "decimal scale" is used to indicate that each digit represents a multiple of a power of 10 (decem), the exponent of the power increasing by 1 each time we move one place to the left. Thus a_0 stands for $a_0 10^0$, a_1 for $a_1 10^1$, a_2 for $a_2 10^2$, etc., each digit for a multiple of a power of 10 with a zero or a positive integral exponent. The use of negative integral exponents enables us to include also "decimal fractions" in our representation. For instance, 45.73, read in the decimal scale, is equivalent to $4 \cdot 10^1 + 5 \cdot 10^0 + 7 \cdot 10^{-1} + 3 \cdot 10^{-2}$. We shall confine our discussion chiefly to natural numbers.

The use of 10 as a basis for a scale of notation is an arbitrary convention; it is a convenient and perhaps natural one, but certainly not necessary. Through long usage we have become so accustomed to the decimal scale, that it is difficult for us to think

of a number apart from its representation in this scale. Our number *words* presuppose its use. One of our present objects is to become acquainted with other scales; the basis may be greater than 10 (major scales), or less than 10 (minor scales). On the history and uses of such scales, the reader will find interesting particulars in several books to which we have referred before.[1]

If we were to use the "septimal scale," every digit would represent the corresponding multiple of a power of 7 (septem). The symbol 43,562 would then stand for $4 \cdot 7^4 + 3 \cdot 7^3 + 5 \cdot 7^2 + 6 \cdot 7 + 2$. It is clear that with this notation we would need only the digits 0, 1, 2, 3, 4, 5, 6. For 7 would be represented by 10, 8 by 11, 9 by 12, 10 by 13, 11 by 14, 12 by 15, 13 by 16, 14 by 20 and so on; it is therefore more economical in the use of digits. Moreover, if this scale were used, it would be very convenient to multiply a number by 7. It should be clear, for instance, that $7 \times 43{,}562$ would be represented by 435,620. Equally simple would be division by 7, since it would merely require that the "septimal point" be moved one place to the left; for instance $43{,}562 \div 7$ would be represented by 4,356.2. Multiplication and division by other numbers would not be very different in one scale from what it is in the other.

Let us try a few examples in the septimal scale. For the product of 5,436 by 5 we would find 40,152. In detail: 5 times 6 is thirty, which is $4 \times 7 + 2$, so that we put down 2 and carry 4; 5 times 3 is fifteen, the 4 that we carried gives nineteen, or $2 \times 7 + 5$, so that we put down 5 and carry 2; 5×4 equals twenty, with the 2 which was carried we obtain twenty-two, which gives 1 and 3 to carry; $5 \times 5 + 3$ gives twenty-eight, so that we put down 0 and carry 4. A little practice would soon enable one to use the septimal scale without any more trouble than is experienced by users of the decimal scale. Division would also lose its terrors. A few more complicated examples will be put down so that in verifying them the reader may test his skill. He should remember that no digit above 6 can occur if the septimal notation is used.

Addition

$$\begin{array}{r} 6504 \\ \underline{2315} \\ 12122 \end{array} \qquad\qquad \begin{array}{r} 34562 \\ \underline{65435} \\ 133330 \end{array}$$

[1] See Dantzig, *op. cit.*, pp. 12–18.

Subtraction

$$\begin{array}{r}6504\\ \underline{2315}\\ 4156\end{array}\qquad\qquad \begin{array}{r}54013\\ \underline{25645}\\ 25035\end{array}$$

Multiplication

$$\begin{array}{r}6504\\ \underline{23}\\ 26115\\ \underline{16311}\\ 222225\end{array}\qquad\qquad \begin{array}{r}46132\\ \underline{536}\\ 412155\\ 204426\\ \underline{333023}\\ 36062045\end{array}$$

Division

```
      22364                    1423003
  12)302031             243)416053210
     24                     243
     32                     1430
     24                     1335
      50                      625
      36                      516
      113                     1063
      105                     1062
        51                       1210
        51                       1062
         0                        115
```

Having acquired a little skill [1] in this new arithmetic we can now venture to carry the two scales side by side and to transpose from one key to the other. We shall put (d) after a number to indicate that it is to be read in the decimal notation and (s) to show that it is to be understood as written in the septimal scale. How can we pass from a number symbol $N(d)$ to a number symbol $M(s)$ and vice-versa, if $N(d)$ and $M(s)$ are to represent the same number?

In the first place, it should be clear that if the number [2] $M(s)$ ends

[1] The process of division will be much facilitated, if we begin by making a table of the multiples of the divisor. For the last example worked out above, this table (in the septimal scale) would be as follows: $1 \times 243 = 243$, $2 \times 243 = 516$, $3 \times 243 = 1062$, $4 \times 243 = 1335$, $5 \times 243 = 1611$, $6 \times 243 = 2154$. Nothing more is needed, since the only digits which could be used in the quotient are 0, 1, 2, 3, 4, 5, 6. We verify incidentally that $7 \times 243 = 2430$.

[2] The word "number" is used here and in what follows with the meaning of "number symbol."

with the digit a_0, then $N(d) = M(s) \equiv a_0$, mod. 7, because $M - a_0$ will then end in 0 and is therefore a multiple of 7. Hence the digit a_0 is simply the residue of N, mod. 7; let us suppose then that $N = 7 \cdot N_1 + a_0$. It remains to convert $N_1(d)$ to a number $M_1(s)$. But if the last digit of $M_1(s)$ is a_1, we have again $N_1(d) = M_1(s) \equiv a_1(\text{mod. } 7)$, so that a_1 is the residue of N_1, mod. 7. If we suppose $N_1 = 7 \cdot N_2 + a_1$, we see that the next digit of $M(s)$ is the residue of N_2, mod. 7. It will now be an easy matter to change $N(d)$ to $M(s)$; we will illustrate the process by an example. To write $48{,}927(d)$ in the septimal scale, we divide the number by 7; we find $48{,}927(d) = 7 \cdot 6{,}989(d) + 4$, so that $N_1 = 6{,}989(d)$ and $a_0 = 4$. The next step then is to divide $6{,}989(d)$ by 7; this gives $6{,}989(d) = 7 \cdot 998(d) + 3$, and therefore $48{,}927(d) = 7[7 \cdot 998(d) + 3] + 4 = 7^2 \cdot 998(d) + 7 \cdot 3 + 4$. Continuing by the same method, we find $998(d) = 7 \cdot 142(d) + 4$, and hence that $48{,}927(d) = 7^3 \cdot 142(d) + 7^2 \cdot 4 + 7 \cdot 3 + 4$. Next $142(d) = 7 \cdot 20(d) + 2$, and $48{,}927(d) = 7^4 \cdot 20(d) + 7^3 \cdot 2 + 7^2 \cdot 4 + 7 \cdot 3 + 4$. Finally $20(d) = 7 \cdot 2 + 6$, so that $48{,}927(d) = 7^5 \cdot 2 + 7^4 \cdot 6 + 7^3 \cdot 2 + 7^2 \cdot 4 + 7 \cdot 3 + 4$. Hence $48{,}927(d) = 262{,}434(s)$.

Let us now do it the other way (man muss immer umkehren!) and convert a number $M(s)$ to a number $N(d)$. The last digit, a_0, of $N(d)$ will be the residue of $M(s)$, mod. $10(d)$, i.e. mod. $13(s)$; to determine it we divide then $M(s)$ by $13(s)$, an operation that we can now perform without much trouble. Suppose then $M(s) = 13(s) \cdot M_1(s) + a_0$; to determine the next digit a_1 of $N(d)$, we divide $M_1(s)$ by $13(s)$, and so forth. For example:[1]

$$\begin{aligned}
65{,}034(s) &= 13(s) \cdot 4{,}464(s) + 6 = 10(d) \cdot 4{,}464(s) + 6;\\
4{,}464(s) &= 13(s) \cdot 320(s) + 4 = 10(d) \cdot 320(s) + 4;\\
320(s) &= 13(s) \cdot 22(s) + 1 = 10 \cdot 22(s) + 1;\\
22(s) &= 13(s) \cdot 1 + 6 = 10 \cdot 1 + 6.
\end{aligned}$$

Therefore

$$65{,}034(s) = 10^4 \cdot 1 + 10^3 \cdot 6 + 10^2 \cdot 1 + 10 \cdot 4 + 6 = 16{,}146(d).$$

This result can also be obtained, indeed much more readily for people still addicted to the decimal system in preference to the septimal, by simply writing

[1] As observed in the footnote on page 197 the division by $13(s)$ is facilitated if we make first a table of the first six multiples of $13(s)$; this table is as follows, everything being written in the septimal scale: $2 \times 13 = 26$, $3 \times 13 = 42$, $4 \times 13 = 55$, $5 \times 13 = 101$, $6 \times 13 = 114$.

$$
\begin{aligned}
65{,}034 &= 4 + 3 \cdot 7 + 0 \cdot 7^2 + 5 \cdot 7^3 + 6 \cdot 7^4 \\
&= 4 + 21 + 5 \cdot 343 + 6 \cdot 2{,}401 \\
&= 25 + 1{,}715 + 14{,}406 = 16{,}146.
\end{aligned}
$$

Of particular interest among the various notations in which numbers can be written are the duodecimal scale, in which 12 is the base (a major scale), and the binary scale, in which 2 is the base, the most minor of scales. If we use the former we need two extra digits, one to designate *ten* and one to designate *eleven;* let us take for these a and b respectively, and let us indicate by (t) placed behind a number symbol that twelve is used as the base. We would then find that $35(d) = 2b(t)$; $593(d) = 49 \cdot 12 + 5(d) = 4 \cdot 12^2 + 1 \cdot 12 + 5(d) = 415(t)$; $7{,}894(d) = 657 \cdot 12(d) + a = 54 \cdot 12^2 + 9 \cdot 12(d) + a = 4 \cdot 12^3 + 6 \cdot 12^2 + 9 \cdot 12(d) + a = 4{,}69a(t)$, etc. In this notation the test for division by 3 is simpler than if the decimal notation is used. For, since a natural number N, whose digits, written in this scale, are $a_k, \ldots a_2, a_1, a_0$, has the value $a_k \cdot 12^k + a_{k-1} \cdot 12^{k-1} + \cdots + a_2 \cdot 12^2 + a_1 \cdot 12 + a_0$, where 12 stands for $12(d)$ it follows that N is exactly divisible by 3 if and only if its last digit, a_0, is divisible by 3, i.e. if it is 0, 3, 6 or 9.

In the binary scale only the digits 1 and 0 are used; for instance we would have the symbols 11, 10101, 111011 for the numbers 3, 21, 59 of the decimal notation respectively. While we need only two different symbols for digits, a good many of each kind are required. The common fraction $\frac{1}{3}$ would take the form of an unending binary fraction, viz. .01010101 . . .; for

$$
\begin{aligned}
\frac{1}{3} = \frac{1}{4} \cdot \frac{4}{3} &= \frac{1}{4}\left(1 + \frac{1}{3}\right) = \frac{1}{4} + \frac{1}{4^2}\left(1 + \frac{1}{3}\right) = \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3}\left(1 + \frac{1}{3}\right) \\
&= \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^6 + \left(\frac{1}{2}\right)^6 \cdot \frac{1}{3} = .010101 + \left(\frac{1}{2}\right)^6 \cdot \frac{1}{3} \\
&= .010101010101 + \left(\frac{1}{2}\right)^{12} \cdot \frac{1}{3},
\end{aligned}
$$

and so forth, the remainder becoming steadily smaller and falling below every positive number taken arbitrarily small.

A few exercises will be useful before we go farther afield.

95. Playing in various keys.

1. Reduce the following numbers from the decimal to the septimal scale: 305, 9843, 6057, 11263.

2. Reduce from the septimal to the decimal scale: 435,600, 12436.

3. Verify the following operations on numbers written in septimal notation:

$$564 + 325 = 1222; \qquad 1042 - 653 = 56;$$
$$343 \times 56 = 30164; \qquad 43032 \div 52 = 561.$$

4. Develop the common fractions $\frac{1}{4}$, $\frac{1}{5}$ and $\frac{2}{7}$ into binary fractions ($\frac{1}{4} = .11$, $\frac{1}{5} = .00110011\ldots$, $\frac{2}{7} = .01001001001\ldots$)

5. Change $433(s)$ from the septimal to the binary scale.

6. Carry out in the binary scale:

$$11{,}011 + 10{,}001; \qquad 1{,}111{,}011 - 100{,}011;$$
$$111 \times 101; \qquad 111{,}001 \div 101.$$

7. Change from the decimal to the duodecimal scale:

$$8{,}754; \quad 20{,}736; \quad 510{,}432.$$

8. Carry out in the duodecimal scale:

$$47a + b56; \quad b783 - 4a0b; \quad 37a \times 2b; \quad b930 \div 95.$$

9. Change $58a7(t)$ from the duodecimal scale to the septimal scale.

10. Change 1,111,010 from the binary scale to the duodecimal scale.

11. Write the following numbers in the duodecimal notation: $78584(s)$; $3692(s)$.

12. (a) If a person's annual salary were $\$2940(t)$, what would be his monthly salary in the duodecimal scale?

(b) If the length of a bar were $11(t)$ feet what would be its length in inches?

13. Show that if and only if the sum of the digits of a number written in the septimal notation is divisible by 3, then the number is itself divisible by 3.

14. Prove that the test for divisibility by 5 of a number written in the septimal scale is as follows. Form a number N by adding together the digits in the 1st, 5th, 9th places, etc., twice the digits in the 2nd, 6th, 10th places, etc., four times the digits in the 3rd, 7th, 11th places, etc., and three times the digits in the 4th, 8th, 12th places, etc. If and only if N is divisible by 5, then the given number is divisible by 5.

15. Develop a test for divisibility by 3 of numbers written in the binary scale; also a test for divisibility by 5.

16. Explain why the test for divisibility by 5 should be the same for numbers written in the binary scale as for numbers written in the septimal scale.

17. Construct a table for the multiplication of the numbers from 1 to 7 in the septimal scale; also for the numbers from 1 to 11 in the scale of which 11 is the base.

18. Prove that if we use a scale whose base is an arbitrary prime

number p, then 1 and $p - 1$ are the only numbers less than p whose square will end in 1.

96. A master scale. Every one doubtless remembers from his school days (compare also *90*, 2) that a number is divisible by 9 if, and only if, when written in the decimal notation, the sum of its digits is divisible by 9. It is an easy matter to show that this test for divisibility by 9 in the decimal notation, is a special case of the following general theorem:

Theorem XLII. A number is divisible by $p - 1$ if and only if, when represented in the scale of base p, the sum of its digits is divisible by $p - 1$.

Proof. Let

$$(9.1)\quad N = a_k p^k + a_{k-1} p^{k-1} + \cdots + a_2 p^2 + a_1 p + a_0,$$

where k is an arbitrary positive integer, and $a_0, a_1, a_2, \ldots a_{k-1}, a_k$ are non-negative integers, *less than* p; then N will be represented in the p-adic scale by the symbol $a_k a_{k-1} \ldots a_0$, consisting of the digits $a_k, a_{k-1}, \ldots a_1, a_0$. Now $p \equiv 1(\text{mod. } p - 1)$; therefore (see Theorem XXXV) $p^2 \equiv 1$, $p^3 \equiv 1$, $p^4 \equiv 1, \ldots p^k \equiv 1$, all mod. $p - 1$. From these facts, it follows by Theorem XXXIV that $a_1 p \equiv a_1$, $a_2 p^2 \equiv a_2, \ldots a_k p^k \equiv a_k(\text{mod. } p - 1)$, and hence by Theorem XXXV, that $N \equiv a_k + a_{k-1} + \cdots + a_1 + a_0(\text{mod. } p - 1)$. It is now evident that N is divisible by $p - 1$, if and only if $a_k + a_{k-1} + \cdots + a_1 + a_0 \equiv 0(\text{mod. } p - 1)$, i.e. if the sum of the digits of N is divisible by $p - 1$.

A similar procedure will enable the reader to prove a theorem for the divisibility by $p + 1$ of a number written in the p-adic scale; as a special case of this theorem he will then obtain Theorem XXXII (see also *99*, 4).

It should be clear[1] that, the terminal digit a_0 of a number N, written in the p-adic scale, is the residue of N, mod. p; i.e. $a_0 \equiv N$, mod. p. Bearing this fact in mind, we will be able to interpret the following observations concerning the residues of certain numbers as statements concerning their terminal digits, when written in the p-adic scale.

Theorem XLIII. The products by $p - 1$ of the numbers 1, 2, 3, . . . $p - 1$ have residues, mod. p, equal respectively to $p - 1$, $p - 2, \ldots 2, 1$.

[1] See also page 198.

Proof. Let a represent any one of the numbers $1, 2, \ldots p-1$. We have then to show that the residue, mod. p, of $a(p-1)$ is equal to $p-a$; i.e. since $p-a<p$, that $a(p-1) \equiv p-a$, mod. p. If this congruence be written in the form

$$ap - a - p + a \equiv 0, \text{ mod. } p,$$

its validity becomes apparent.

Remark. A special case of this theorem is the well-known fact that the products by 9 of the numbers $1, 2, \ldots 9$ end in $9, 8, \ldots 1$ respectively, when written in the decimal scale.

Corollary 1. Among the numbers $1, 2, \ldots p-1$, there is only one, viz. $p-1$, whose product by $p-1$ is congruent to 1, mod. p.

We can go a step further. For, since $a(p-1) \equiv p-a$, mod. p, it follows from Theorem XXXIV that

$$a(p-1)^2 \equiv (p-a)(p-1) = p^2 - ap - p + a \equiv a, \text{ mod. } p,$$

so that the products of the numbers $1, 2, \ldots p-1$ by $(p-1)^2$ have residues, mod. p, equal to $1, 2, \ldots p-1$ respectively. But $p-1 \equiv -1$, mod. p; hence we can say that among the numbers $1, 2, \ldots p-1$, there is only one, viz. $p-1$, whose product by $(p-1)^2$ is congruent to -1, mod. p; i.e. $(p-1)^3$ is the only one among these products which ends in $p-1$, when written in the p-adic scale. Continuing in this way, we obtain the following result:

Corollary 2. Among the numbers $1, 2, \ldots p-1$, there is only one, viz. $p-1$, whose product by $(p-1)^k$ is congruent to 1, mod. p, when k is odd; and only one, viz. $p-1$, whose product by $(p-1)^k$ is congruent to -1, when k is even.

We observe, for future use, that if $a \equiv b$, mod. p, and a and b have a common factor m, which is relatively prime to p, say $a = ma_1$ and $b = mb_1$, then $a_1 \equiv b_1$, mod. p (compare *99*, 6).

97. An old chestnut in a new shell. Some years ago a puzzle, published in one of the weekly journals, agitated a great many people who tried to solve it. The question was this: A pile of cocoanuts is discovered by a monkey; he throws away the first one and carries off exactly one fifth of what is left. A second monkey imitates the first, then a third, a fourth and a fifth; each time, after having thrown one nut away, they can take exactly one fifth of what is left. The pile which finally remains is exactly

divisible by 5. What is the least number of nuts that can have been in the original pile? Although it is a pity to spoil the fun for future generations, we will show how easily this question can be answered by use of the tools we have just acquired. Let us suppose that, instead of five, there were p monkeys, and that each one instead of taking a fifth part, took a pth part. Let us then denote the original number of nuts by N, and let us think of N as represented in the scale of base p, as in (9.1); furthermore let us denote by N_1 the number of nuts left after the first monkey had taken his part, by N_2 the number left after the second had finished his work, etc. Then we shall consider the following slightly modified "monkey and cocoanut" problem:

To determine the digits $a_0, a_1, \ldots a_k$ of a number N written in the p-adic scale such that $N \equiv 1 (\text{mod. } p)$, $N_1 \equiv 1 (\text{mod. } p)$, $N_2 \equiv 1 (\text{mod. } p)$; $\ldots N_k \equiv 1 (\text{mod. } p)$.[1]

Solution. From the first of these congruences, we conclude that $a_0 = 1$. Furthermore N_1 is obtained from N by subtracting 1 and then diminishing the remainder by its pth part; i.e.

$$N_1 = N - 1 - \frac{N-1}{p} = \frac{p(N-1)-(N-1)}{p} = \frac{(N-1)(p-1)}{p}.$$

It follows from (9.1) that

$$N_1 = (a_k p^{k-1} + a_{k-1} p^{k-2} + \cdots + a_2 p + a_1)(p-1),$$

so that $N_1 \equiv a_1(p-1)$, mod. p. Since the conditions of the problem require that $N_1 \equiv 1$, mod. p, we conclude from Theorem XXXIII that we must choose a_1 so that $a_1(p-1) \equiv 1$, mod. p. Since $a_1 < p$, Cor. 1 of Theorem XLIII becomes applicable; by means of it, we are led to the conclusion that $a_1 \equiv p - 1$, and that

$$(9.2) \quad N_1 \equiv (a_k p^{k-1} + a_{k-1} p^{k-2} + \cdots + a_2 p)(p-1) + (p-1)^2.$$

We have now determined two of the digits of N. To find the next one, a_2, we proceed very much in the same way; we imitate the previous procedure, just as the monkeys did. We find then, since $N_2 = \dfrac{(N_1 - 1)(p-1)}{p}$, that we can write, by use of (9.2)

$$N_2 = (a_k p^{k-2} + a_{k-1} p^{k-3} + \cdots + a_3 p + a_2)(p-1)^2 + (p-2)(p-1).$$

[1] It is not stipulated that the number N is to have exactly $k+1$ digits; we are only asked to determine the last $k+1$ of its digits.

Hence, $N_2 \equiv a_2(p-1)^2 + (p-2)(p-1)$, mod. p, from which it follows, by use of Theorem XXXIII that the condition $N_2 \equiv 1$, mod. p, is equivalent to the requirement that

$$a_2(p-1)^2 + (p-2)(p-1) \equiv 1, \text{ mod. } p.$$

Since $1 - (p-2)(p-1) = -p^2 + 3p - 1 \equiv -1$, mod. p, this requirement reduces to $a_2(p-1)^2 \equiv -1$, mod. p. We can now make use of Cor. 2 of Theorem XLIII to reach the conclusion that $a_2 = p - 1$. Using this value for a_2, we find that

$$N_2 = (a_k p^{k-2} + a_{k-1}p^{k-3} + \cdots + a_3 p)(p-1)^2 + (p-1)^3 + (p-2)(p-1),$$

i.e.

$$N_2 = (a_k p^{k-2} + a_{k-1}p^{k-3} + \cdots + a_3 p)(p-1)^2 + p(p-1)^2 - p + 1.$$

In this way we could continue; we would find that

$$a_3 = a_4 = \cdots = a_{10} = p - 1.$$

To avoid wearisome, endless, monkey-like and moreover futile[1] imitation, we turn to the principle of mathematical induction (compare p. 91) to complete the discussion. Suppose we have proved that the formulas

(9.3)

$$N_i = (a_k p^{k-i} + a_{k-1}p^{k-i-1} + \cdots + a_{i+1}p)(p-1)^i + p(p-1)^i - p + 1,$$

and $$a_i = p - 1$$

hold for $i = 1, 2, \ldots j - 1$.[2] Then we know that

$$N_{j-1} = (a_k p^{k-j+1} + a_{k-1}p^{k-j} + \cdots + a_j p)(p-1)^{j-1} + p(p-1)^{j-1} - p + 1,$$

and hence that

(9.31)

$$N_j = \frac{(N_{j-1} - 1)(p-1)}{p} = (a_k p^{k-j} + \cdots + a_j)(p-1)^j + (p-1)^j - (p-1),$$

so that

$$N_j \equiv a_j(p-1)^j + (p-1)^j - p + 1, \text{ mod. } p.$$

The condition $N_j \equiv 1$, mod. p, is then equivalent to the condition $a_j(p-1)^j + (p-1)^j \equiv 0$, mod. p. Since now $p - 1 \equiv -1$, mod. p, it follows from Theorem XXXV that $(p-1)^j \equiv (-1)^j$,

[1] Futile, because the value of k is not specified.
[2] These formulas have actually been proved for $i = 1, 2$.

mod. p. Hence, the condition that $N_j \equiv 1$, mod. p, has been reduced to the condition $a_j(p-1)^j \equiv -(p-1)^j \equiv -(-1)^j = (-1)^{j+1} = 1$ or -1, according as j is odd or even. By means of Cor. 2, we conclude now that $a_j = p - 1$, and also, by use of (9.31), that

$$\begin{aligned} N_j &= (a_k p^{k-j} + \cdots + a_{j+1}p)(p-1)^j + (p-1)^{j+1} + (p-1)^j - p + 1 \\ &= (a_k p^{k-j} + \cdots + a_{j+1}p)(p-1)^j + p(p-1)^j - p + 1. \end{aligned}$$

Thus we see that (9.3) must hold for $i = j$, if it holds for $i = j - 1$.

Since we have already shown that (9.3) holds for $i = 1, 2$, the principle of mathematical induction justifies the conclusion that $a = p - 1$, for $i = 1, 2, \ldots k$. The number N which we set out to determine must then, if written in the p-adic scale, terminate with the digits: $p-1, p-1, \ldots p-1, 1$; any number which ends with these digits has the property required in the statement of the problem, no matter what the preceding digits are. Consequently, the least number of this kind is

$$(p-1)p^k + (p-1)p^{k-1} + \cdots + (p-1)p^2 + (p-1)p + 1.$$

98. Crab fashion. In preparation for the discussion of some further interesting facts about numbers, we recall the observation, made on page 57, that for any natural number k, and for any two integers a and b, $a^k - b^k$ is divisible by $a - b$, since

$$a^k - b^k = (a-b)(a^{k-1} + a^{k-2}b + \cdots + ab^{k-2} + b^{k-1}).$$

We add to this two further remarks, viz.,

1. if k is *odd*,

$$a^k + b^k = (a+b)(a^{k-1} - a^{k-2}b + \cdots - ab^{k-2} + b^{k-1}),$$

so that $a^k + b^k$ is then divisible by $a + b$; and

2. if k is *even*,

$$a^k - b^k = (a+b)(a^{k-1} - a^{k-2}b + \cdots + ab^{k-2} - b^{k-1}),$$

so that $a^k - b^k$ is then divisible, not only by $a - b$, but also by $a + b$.[1]

[1] These two remarks can readily be justified by a little multiplication and the use of mathematical induction; we have only to notice that

$$a^{k+2} + b^{k+2} = (a^k + b^k)(a^2 + b^2) - (a^{k-2} + b^{k-2})a^2b^2,$$

and that

$$a^{k+2} - b^{k+2} = (a^k - b^k)(a^2 + b^2) - (a^{k-2} - b^{k-2})a^2b^2,$$

so that the statements 1 and 2 are true for any odd (even) number k if they are true for all smaller odd (even) numbers. The reader who prefers not to take this trouble can

Let us now once more consider an arbitrary number N, written in the p-adic scale, as in (9.1),

$$N = a_k p^k + a_{k-1}p^{k-1} + \cdots + a_2 p^2 + a_1 p + a_0;$$

and let us consider along with it the number M obtained by writing N in crab fashion [1] so that

$$M = a_k + a_{k-1}p + \cdots + a_2 p^{k-2} + a_1 p^{k-1} + a_0 p^k.$$

What can we say about the numbers $N - M$ and $N + M$? (Let us suppose $N \geqq M$; there is no essential change in the reasoning if $N < M$.) We find that

$$N - M = a_k(p^k - 1) + a_{k-1}(p^{k-1} - p) + \cdots + a_2(p^2 - p^{k-2}) + a_1(p - p^{k-1}) + a_0(1 - p^k).$$

If k is even, we can put $k = 2m$, where m is a natural number, and we shall find in the center of the expression for $N - M$ (notice that the exponent in the first term in each pair of parentheses is equal to the subscript of the coefficient on the outside, and that the sum of the exponents of the two terms in each pair of parentheses is always equal to k) the terms

$$a_{m+1}(p^{m+1} - p^{m-1}) + a_m(p^m - p^m) + a_{m-1}(p^{m-1} - p^{m+1}).$$

If k is odd, we can put $k = 2m + 1$, and the center will consist of the two terms

$$a_{m+1}(p^{m+1} - p^m) + a_m(p^m - p^{m+1}).$$

Therefore we can say:

(9.4) If $k = 2m$,

$$\begin{aligned} N - M = {} & a_{2m}(p^{2m} - 1) + a_{2m-1}(p^{2m-2} - 1)p + \cdots \\ & + a_{m+1}(p^2 - 1)p^{m-1} + a_{m-1}(1 - p^2)p^{m-1} + \cdots \\ & + a_2(1 - p^{2m-4})p^2 + a_1(1 - p^{2m-2})p + a_0(1 - p^{2m}); \end{aligned}$$

and

(9.41) if $k = 2m + 1$,

accept the statement upon the authority of one of the numerous writers on elementary algebra; he will do so at his own risk.

The observation of page 57, referred to above, can also be interpreted as stating that the geometrical progression whose first term is a^{k-1} (or b^{k-1}) and whose ratio is $\frac{b}{a}\left(\text{or } \frac{a}{b}\right)$ has a sum equal to $\frac{a^k - b^k}{a - b}$. This statement can easily be seen to include the usual formula for the sum of a geometrical progression given in algebra textbooks.

[1] The locomotive habits of the crab are certainly known to every reader; we shall call M the "crab-image" of N.

$$N - M = a_{2m+1}(p^{2m+1} - 1) + a_{2m}(p^{2m-1} - 1)p + \cdots$$
$$+ a_{m+1}(p - 1)p^m + a_m(1 - p)p^m + \cdots$$
$$+ a_2(1 - p^{2m-3})p^2 + a_1(1 - p^{2m-1})p + a_0(1 - p^{2m+1}).$$

In the first case, the expression inside each pair of parentheses is the difference of two equal *even* powers of p and 1 (remember that every power of 1 equals 1), so that in virtue of remark 2 above, it is divisible by $p - 1$ and by $p + 1$. In the second case, the expressions inside all the pairs of parentheses are the differences of two equal *odd* powers of p and 1 and are therefore divisible by $p - 1$, as we have known since page 57. We can therefore conclude from (9.4) and (9.41), that *if k is even the difference $N - M$ is divisible by $p - 1$ and by $p + 1$, while if k is odd the difference $N - M$ is divisible by $p - 1$.* The reader should have no difficulty in deriving for $N + M$ the following expressions, analogous to (9.4) and (9.41):

(9.5) If $k = 2m$,

$$N + M = a_{2m}(p^{2m} + 1) + \cdots + a_{m+1}(p^2 + 1)p^{m-1} + 2a_m p^m$$
$$+ a_{m-1}(1 + p^2)p^{m-1} + \cdots + a_0(1 + p^{2m}),$$

and

(9.51) if $k = 2m + 1$,

$$N + M = a_{2m+1}(p^{2m+1} + 1) + \cdots + a_{m+1}(p + 1)p^m$$
$$+ a_m(1 + p)p^m + \cdots + a_0(1 + p^{2m+1}).$$

From the former of these formulae we can not make any general deduction; but from the latter we conclude by means of remark 1, that $N + M$ is divisible by $p + 1$. The discussion has yielded the interesting result stated in the following theorem; in reading it the reader must remember that the number of digits in N as given by (9.1) is equal to $k + 1$, so that it is even if k is odd and odd if k is even.

Theorem XLIV. If the number of digits of a number N written in the p-adic scale is *odd*, then the difference between it and its crab-image is divisible by $p - 1$ and by $p + 1$; if the number of digits is *even*, then the difference between it and its crab-image is divisible by $p - 1$, while the sum of the two is divisible by $p + 1$.

For the special case in which the decimal scale is used, the results stated in this theorem can readily be tested by the reader on numerical examples; if he has gained some skill in performing

in major and minor scales, he will not have to limit himself however to this special case. No matter in what scale the numbers are interpreted, provided only that the base p exceeds by at least one the largest digit occurring in the number, it must be true, that a difference like $57834 - 43875$ is divisible by $p - 1$ and by $p + 1$; that $965325 - 523569$ is divisible by $p - 1$ and $965325 + 523569$ by $p + 1$.[1]

The following simple little game is based on the first part of this theorem: A asks B to write down any number of 3 digits (in the decimal scale). Then he writes under it the crab-image of the number B has written and asks B to subtract the smaller of the two numbers from the larger and to divide the difference by 99; the result will always be the difference between the first and last digits of his number – and is therefore predictable. If B writes down 479, A forces him to obtain the result 5 by writing down 974 and asking B to subtract the first from the second and to divide the difference by 99. The reader who has followed the argument on the preceding page should have no difficulty in accounting for the result. He will also be able to devise other parlor tricks of similar character.

99. Games and Puzzles.

1. Prove that a number which is represented in the decimal scale by a symbol of the type *abcabc* is always divisible by 7, by 11 and by 13.

2. Show that the residue, mod. 8, of the square of any odd number is equal to 1.

3. Show that the residue, mod. 9, of the cube of any integral number is equal to 1, or 0, or -1.

4. Develop a test for divisibility by $p + 1$ of numbers written in the p-adic scale; apply to various special cases (compare p. 201).

5. Prove that if p is *even* none of the products of the numbers 1, 2, . . . $p - 1$ by $p - 2$, or by $p - 4$, etc., has the residue 1, mod. p.

6. Show that if $a = ka_1$, $b = kb_1$, $a \equiv b$, mod. p, and k is relatively prime to p, then $a_1 \equiv b_1$, mod. p (see final paragraph of *96*).

7. Prove that if N is written in the decimal scale and if M is its crab-image, then $N + M$ is divisible by 9 if and only if N is.

8. Extend the proposition of 7 to an arbitrary scale.

[1] Readers who are interested in aspects of the theory of numbers illustrated above will find material to their taste in E. Lucas, *L'Arithmétique Amusante;* see also Guttman, "On Cyclic Numbers," *American Mathematical Monthly*, vol. 41, 1934, p. 159 and the references at the end of *102*.

9. Prove that under the hypothesis of 7 and with the further assumption that N has an odd number of digits, $N + M$ is divisible by 11 if and only if N is.

10. Prove that under the hypothesis of 7 and the further assumption that N has an even number of digits, $N - M$ is divisible by 11 only if N is.

11. Prove that whenever $a \not\equiv 0 (\text{mod. } 3)$ *and* $b \not\equiv 0 (\text{mod. } 3)$, one of the numbers $a + b$ and $a - b$ is congruent to 0, mod. 3; the conclusion is obviously valid if $a \equiv b \equiv 0 (\text{mod. } 3)$. Show moreover that neither $a + b$ nor $a - b$ are congruent to 0, unless *either* both a and b are incongruent to 0 or both are congruent to $0 (\text{mod. } 3)$.

12. Prove that if a and b are two numbers of equal parity then $ab(a + b)(a - b)$ is always divisible by 24.

13. Show that if N is any number of n digits written in the decimal scale, and M the number obtained from N by merely interchanging the first and last digits, then the quotient of their difference by the difference between the first and last digits will always consist of $n - 1$ nines.

14. Show that, if the sum of any four-digit number written in the decimal scale and its crab-image be divided by 11, the quotient will always be equal to 91 times the sum of the first and last digits increased by 10 times the sum of the middle digits.

15. Show that, if the difference between any four-digit number and its crab-image be divided by 9, the quotient will always be equal to 111 times the difference of the first and fourth digits, increased by 10 times the difference of the second and third digits.

16. Extend the statements in 14 and 15 to four-digit numbers in a p-adic scale.

17. Show that any equality between number expressions involving additions, subtractions and multiplications must also hold if all the numbers are replaced by their residues, with respect to an arbitrary modulus.

Note: It is upon this fact that the "casting out of nines" procedure was based, used by our grandfathers to test their calculations. To test whether indeed $\dfrac{17 \times 56}{28} = 34$, they would write it in the form $17 \times 56 = 34 \times 28$; then they would cast out nines, i.e. in our present terminology determine the residues, mod. 9, of the numbers on each side of the supposed equality, and test whether the residues are equal. In this instance we find

$$17 \times 56 \equiv 8 \times 2 = 16 \equiv 7 (\text{mod. } 9); \quad \text{and } 34 \times 28 \equiv 7 \times 1 = 7 (\text{mod. } 9).$$

Compare, e.g., Cajori, *History of Mathematics*, pp. 91–106.

18. Prove that $p - 1$ and p are always relatively prime.

100. The divisors of a number. If the number p is a prime, it has no divisors, except itself and 1 (see p. 8). Which are the divisors of p^2? Since p^2 has no divisors which p does not have, it should be clear that its only factors are 1, p and p^2. Similarly the only divisors of p^3 are 1, p, p^2 and p^3; and, in general, the only divisors of p^n, for any natural number n, are 1, p, p^2, . . . p^{n-1}, p^n, there being in all $n + 1$ divisors. Consequently, the *sum* of the divisors of p^n is equal to $1 + p + p^2 + \cdots + p^n$. If we recall the opening remark of *98*, we know that this sum is equal to $\frac{p^{n+1} - 1}{p - 1}$, i.e. the sum of the divisors of p^n, where p is a prime, is equal to $\frac{p^{n+1} - 1}{p - 1}$. For example, consider the number $16 = 2^4$, for which $p = 2$ and $n = 4$. Its divisors are 1, 2, 4, 8, 16; their sum is 31. On the other hand $\frac{p^{n+1} - 1}{p - 1}$ becomes in this case $\frac{2^5 - 1}{2 - 1}$, which is also equal to 31.

Suppose now that q is another prime number. Which numbers will then be divisors of p^nq? Clearly all the $n + 1$ numbers which are divisors of p^n are among them; but furthermore the products of these numbers by q are also divisors of p^nq. And, no other number can be, because, in order to be a divisor of p^nq, a number must have for its factors some or all of the factors of p^n and of q. Continuing in this way, we reach the conclusion that the factors of p^nq^m are the numbers of all the sets 1, p, p^2, . . . p^{n-1}, p^n; q, pq, p^2q, . . . $p^{n-1}q$, p^nq; q^2, pq^2, p^2q^2, . . . $p^{n-1}q^2$, p^nq^2; . . .; q^{m-1}, pq^{m-1}, p^2q^{m-1}, . . . $p^{n-1}q^{m-1}$, p^nq^{m-1}; q^m, pq^m, p^2q^m, . . . $p^{n-1}q^m$, p^nq^m. There are thus $m + 1$ sets of $n + 1$ divisors each, so that the total number of the divisors is $(n + 1)(m + 1)$. Moreover their *sum* consists of the sum of the terms in the expanded product $(1 + p + p^2 + \cdots + p^{n-1} + p^n)(1 + q + q^2 + \cdots + q^{m-1} + q^m)$; but this is clearly equal to

$$\frac{p^{n+1} - 1}{p - 1} \cdot \frac{q^{m+1} - 1}{q - 1}.$$

For example, the sum of the divisors of $36 = 2^2 \cdot 3^2$ is equal to

$$\frac{(2^3 - 1)(3^3 - 1)}{(2 - 1)(3 - 1)} = \frac{7 \cdot 26}{2} = 91;$$

this is readily verified if we notice that the divisors of 36 are 1, 2, 3, 4, 6, 9, 12, 18, 36.

We are now prepared to prove a general theorem. First we observe that, if a number N is resolved into its prime factors, it takes the form $N = p_1^{k_1} p_2^{k_2} \ldots p_{n-1}^{k_{n-1}} p_n^{k_n}$, where $p_1, p_2, \ldots p_n$ are different prime numbers and $k_1, k_2, \ldots k_n$ are arbitrary natural numbers. Any divisors of N will therefore consist of a product of some factors of $p_1^{k_1}$ (or none, or all), some factors of $p_2^{k_2}$ (or none, or all), . . ., and some factors of $p_n^{k_n}$ (or none, or all); if the "none" applies in all cases we obtain the divisor 1, if the "all" applies in all cases we find the factor N. Therefore the divisors of N are the terms which we get if we expand the product $(1 + p_1 + p_1^2 + \cdots + p_1^{k_1})(1 + p_2 + p_2^2 + \cdots + p_2^{k_2}) \cdots (1 + p_n + p_n^2 + \cdots + p_n^{k_n})$. We have only to remember now some of our earlier remarks to recognize the validity of the following statement:

Theorem XLV. The number of divisors of the number $N = p_1^{k_1} p_2^{k_2} \ldots p_n^{k_n}$, in which $p_1, \ldots p_n$ are different primes and $k_1, \ldots k_n$ are arbitrary natural numbers, is equal to

$$(k_1 + 1)(k_2 + 1) \ldots (k_n + 1);$$

the sum of the divisors is equal to

$$\frac{(p_1^{k_1+1} - 1)(p_2^{k_2+1} - 1) \cdots (p_n^{k_n+1} - 1)}{(p_1 - 1)(p_2 - 1) \cdots (p_n - 1)}.$$

To apply this theorem to the number 1800, we factor it into prime factors, which shows us that $1800 = 2^3 \cdot 3^2 \cdot 5^2$. From this we conclude that 1800 has $(3 + 1)(2 + 1)(2 + 1) = 36$ factors, and that their sum is equal to $\dfrac{(2^4 - 1)(3^3 - 1)(5^3 - 1)}{(2 - 1)(3 - 1)(5 - 1)}$ $= \dfrac{15 \cdot 26 \cdot 124}{2 \cdot 4} = 15 \cdot 13 \cdot 31 = 6045$; the ambitious calculator may verify the result.

101. Perfect numbers. Besides classifying the natural numbers into the two classes of *odd* and *even* numbers, the Pythagoreans classified them with reference to the sum of their factors.[1] A number N, for which the sum of the factors, exclusive of N itself, is less than N was called *defective;* if this sum exceeds N, the number was

[1] See e.g., Cajori, *op. cit.*, p. 68; Sanford, *op. cit.*, p. 330.

called *abundant* (or excessive), and if N = the sum of the factors of N, excluding N itself, then N was called a *perfect* number. For example, since the factors of 8, different from 8 itself, are 1, 2 and 4, whose sum is 7, the number 8 is defective; similarly since $12 = 2^2 \cdot 3$, the sum of all its factors is $\frac{(2^3 - 1)(3^2 - 1)}{2} = 28$ and therefore the sum of the factors, different from 12, is equal to 16 — it is an abundant number. But 6, which has the factors 1, 2 and 3, whose sum is 6, is a perfect number.

The search for perfect numbers reveals some interesting facts. In order to avoid a longer form of statement, we will take the definition in the following equivalent form:

Definition XXXVIII. A number N the sum of whose factors is equal to $2N$ is called a perfect number.

It follows from Theorem XLV that if the number $N = p_1^{k_1}p_2^{k_2} \ldots p_n^{k_n}$ is a perfect number, we must have

$$\begin{aligned}(9.6)\quad (p_1^{k_1+1} - 1)(p_2^{k_2+1} - 1) \ldots (p_n^{k_n+1} - 1)\\ = 2(p_1 - 1)(p_2 - 1) \ldots (p_n - 1)p_1^{k_1}p_2^{k_2} \ldots p_n^{k_n};\end{aligned}$$

and that, whenever this condition is fulfilled, N is a perfect number.

If the general solution of equation (9.6) could be found in the sense in which (1.2) furnishes the general solution of equation (1.1) we would have obtained all perfect numbers. Such a general solution has not been discovered; but Euclid showed that the search for it is intimately linked up with the search for prime numbers. Of this connection we shall now try to get some idea.

A few preliminary reconnoitering expeditions will be useful before an attack is made on the main problem.

Theorem XLVI. No number of the form p^k is a perfect number, if p is a prime.

Proof. If there were a perfect number of the form p^k, in which p is a prime, then p and k would have to satisfy the equation obtained from (9.6) for $n = 1$, viz. dropping the subscript 1:

$$p^{k+1} - 1 = 2(p - 1)p^k.$$

This equation is readily transformed to $p^{k+1} - 2p^k + 1 = 0$, and then to the form $p^k(p - 1) = p^k - 1$. Since now $p^k > p^k - 1$, and since $p - 1 \geqq 1$, it follows (remember *15*, 3, p. 28), that this equation can not be satisfied. Thus the theorem is proved.

Having conquered this first outpost, let us advance a step further.

Theorem XLVII. If $N = p_1^{k_1}p_2^{k_2}$, where p_1 and p_2 are primes, is a perfect number, then p_1 and p_2 can not both be odd.

Proof. If this number N were perfect, we would obtain from (9.6) the condition

$$(p_1^{k_1+1} - 1)(p_2^{k_2+1} - 1) = 2(p_1 - 1)(p_2 - 1)p_1^{k_1}p_2^{k_2}.$$

From *98*, we know that both sides of this equation may be divided by $(p_1 - 1)(p_2 - 1)$ and that the resulting equation is:

(9.7)

$$(p_1^{k_1} + p_1^{k_1-1} + \cdots + p_1 + 1)(p_2^{k_2} + p_2^{k_2-1} + \cdots + p_2 + 1) = 2p_1^{k_1}p_2^{k_2}.$$

If p_1 and p_2 are both odd primes, the right side of this equation has a single factor 2. What about the left side? The first pair of parentheses contains the sum of $k_1 + 1$ odd numbers, the second pair the sum of $k_2 + 1$ odd numbers.

(a) If $k_1 + 1$ and $k_2 + 1$ are both odd, each factor on the left is an odd number,[1] their product is therefore odd and can not have the factor 2 which appears on the right side.

(b) [2] If $k_1 + 1$ is even, the terms in the first pair of parentheses can be grouped, two by two, $(p_1^{k_1} + p_1^{k_1-1}) + (p_1^{k_1-2} + p_1^{k_1-3}) + \cdots + (p_1^3 + p_1^2) + (p_1 + 1)$, so that each group has the factor $p_1 + 1$; the left side has then the factor $p_1 + 1$, and the equation (9.7) takes the form

$$(p_1 + 1)(p_1^{k_1-1} + p_1^{k_1-3} + \cdots + p_1^2 + 1)(p_2^{k_2} + p_2^{k_2-1} + \cdots + p_2 + 1) = 2p_1^{k_1}p_2^{k_2}.$$

Now $p_1 + 1$ is even and > 2; since it must be a factor of the right side and since it does not have a factor p_1 it must have one or more factors p_2; let us suppose therefore $p_1 + 1 = 2p_2^k$, i.e. $p_1 = 2p_2^k - 1$. The equation then becomes

$$(p_1^{k_1-1} + p_1^{k_1-3} + \cdots + p_1^2 + 1)(p_2^{k_2} + p_2^{k_2-1} + \cdots + p_2 + 1) = p_1^{k_1}p_2^{k_2-k}.$$

[1] The reader will easily convince himself of the fact that the sum of an odd number of odd numbers is odd, and the sum of an even number of odd numbers is even; i.e. that the sum of n odd numbers has the same parity as n.

[2] In working through the following argument, the reader will find it useful to recall the reasoning on page 192.

The first factor on the left of this new equation is congruent to 1, mod. p_1, so that it has no factor p_1. Therefore the second factor on the left must be divisible by $p_1^{k_1}$, i.e. by $(2p_2^k - 1)^{k_1}$; and since this second factor is not divisible by p_2, it must be equal to $(2p_2^k - 1)^{k_1}$, so that we must have

$$p_2^{k_2} + p_2^{k_2-1} + \cdots + p_2 + 1 = (2p_2^k - 1)^{k_1}.$$

Now $2p_2^k - 1 \equiv -1$, mod. p_2, and hence $(2p_2^k - 1)^{k_1} \equiv (-1)^{k_1}$, mod. p_2. But $k_1 + 1$ is even, hence k_1 odd; consequently the right side of the last equation is congruent to -1, mod. p_2. On the other hand, the residue of the left side, mod. p_2, is $+1$. We have now arrived at a contradiction, since, if p_2 is odd, $+1$ can not be congruent to -1, mod. p_2. Hence (9.7) can not hold if $k_1 + 1$ is even. In exactly similar fashion we prove that it can not hold if $k_2 + 1$ is even. Since the argument in (b) is entirely independent of the parity of k_2, we have exhausted all the possibilities with respect to k_1 and k_2; there is therefore no escape from the statement made in Theorem XLVII.

What possibilities are left for the perfect numbers? Perfection, even among numbers, is evidently not easy to attain. Since the only prime, which is not odd, is 2, it follows that if N is to be a perfect number of the form $p_1^{k_1}p_2^{k_2}$ in which p_1 and p_2 are primes, then one of these primes must be 2. What about the other?

Let us suppose that $p_1 = 2$, and that $N = 2^{k_1}p_2^{k_2}$ is a perfect number. Then it follows that

$$(2^{k_1+1} - 1)(p_2^{k_2+1} - 1) = 2^{k_1+1}(p_2 - 1)p_2^{k_2},$$

and hence that

$$2^{k_1+1}(p_2^{k_2} - 1) = p_2^{k_2+1} - 1. \tag{9.8}$$

Consequently $p_2^{k_2+1} - 1$ must be divisible by $p_2^{k_2} - 1$. Since however, $p_2^{k_2+1} - 1 = p_2(p_2^{k_2} - 1) + p_2 - 1$, this carries with it that $p_2 - 1$ is divisible by $p_2^{k_2} - 1$, which is possible if and only if $k_2 = 1$. But then (9.8) becomes $p_2^2 - 1 = 2^{k_1+1}(p_2 - 1)$, from which we conclude that $p_2 = 2^{k_1+1} - 1$.

Corollary. The only perfect numbers N of the form $p_1^{k_1}p_2^{k_2}$ where p_1 and p_2 are primes, are the numbers given by the formula $N = 2^{k_1}(2^{k_1+1} - 1)$, in which k_1 has such values that $2^{k_1+1} - 1$ is a prime.

This result is completed in an essential way by two additional facts; the first of these is the following:

Theorem XLVIII. Whenever $2^{k+1} - 1$ is a prime, the number $N = 2^k(2^{k+1} - 1)$ is perfect.

Proof. Under the hypothesis, the factors of N, exclusive of N itself, are: $1, 2, 2^2, \ldots 2^k$ and $m, 2m, 2^2m, \ldots 2^{k-1}m$, in which m is an abbreviation for the prime $2^{k+1} - 1$. The sum of these factors is $1 + 2 + 2^2 + \cdots + 2^k + m(1 + 2 + \cdots + 2^{k-1}) = 2^{k+1} - 1 + m(2^k - 1)$, compare *98*. But $2^{k+1} - 1 + m(2^k - 1) = 2^{k+1} + m \cdot 2^k - (m + 1) = 2^{k+1} + m \cdot 2^k - 2^{k+1} = m \cdot 2^k = N$. Therefore N is a perfect number.

Thus we see that the search for perfect numbers is closely connected with the search for prime numbers of the form $2^{k+1} - 1$. If $k + 1$ is composite, let us say $k + 1 = pq$, then $2^{k+1} - 1 = (2^p)^q - 1$, so that the number $2^{k+1} - 1$ is divisible by $2^p - 1$ (and also by $2^q - 1$) and therefore composite. Consequently $2^{k+1} - 1$ can be a prime only if $k + 1$ is a prime. Unfortunately, the converse is not true; even if $k + 1$ is a prime, $2^{k+1} - 1$ may be composite. The smallest prime which shows this is 11; in fact $2^{11} - 1 = 2047 = 23 \cdot 89$.

Next, it should be observed that the foregoing discussion has not shown that *every* perfect number is of the form 2^km, where m is a prime of the form $2^{k+1} - 1$, but only that perfect numbers of the form $p_1^{k_1}p_2^{k_2}$ are of this type. There remains the possibility that there are perfect numbers of the form $p_1^{k_1}p_2^{k_2} \ldots p_n^{k_n}$, discussed in Theorem XLV in which $p_1, p_2, \ldots p_n$ are all *odd* primes and $n > 2$.

Whether or not this is so, is still an open question; the result stated in Theorems XLVI–XLVIII does not touch it and it is doubtful whether the method used in the proof of this theorem can be made to serve in dealing with this more difficult problem. It is however possible to show that *every even* perfect number is of the form 2^km; we conclude with a proof of this fact.

Theorem XLIX. Every even perfect number N has the form 2^km, where m is a prime of the form $2^{k+1} - 1$.

Proof. Suppose $N = 2^km$, where $k > 0$ and m is *odd*, is a perfect number. Let the divisors of m, exclusive of m, be $1, m_1, m_2, \ldots m_M$ and put $1 + m_1 + m_2 + \cdots + m_M = M_1$. Then the sum of *all* the divisors of N is equal to

$$(1 + 2 + \cdots + 2^k)(1 + m_1 + \cdots + m_M + m) = (2^{k+1} - 1)(M_1 + m).$$

Therefore, since N is a perfect number, we must have $2N = 2^{k+1}m = (2^{k+1} - 1)(M_1 + m) = 2^{k+1}M_1 + 2^{k+1}m - (M_1 + m)$; therefore

$$(9.9) \qquad m = (2^{k+1} - 1)M_1.$$

It follows that M_1 is a divisor of m; but it can not be m itself, since otherwise $2^{k+1} - 1 = 1$, which would mean that $k = 0$, contrary to hypothesis. Therefore M_1 must be equal to one of the numbers $1, m_1, m_2, \ldots m_M$ of which it is the sum. This is possible only if $M_1 = 1$ and $m_1 = m_2 = \cdots = m_M = 0$. But this means that the only divisors of m are 1 and m, so that m is a prime number; moreover it follows then from (9.9) that $m = 2^{k+1} - 1$. We conclude that $N = 2^k(2^{k+1} - 1)$, as was asserted in the statement of our theorem.[1]

102. To redeem a promise. A concept related to the idea of perfect numbers is that of a pair of *amicable numbers*, i.e. a pair of numbers N and M such that either is equal to the sum of the divisors of the others, not counting the numbers themselves among their divisors.

An example of a pair of amicable numbers is furnished by 220 and 284; for the factors of the first of these numbers are 1, 2, 4, 5, 10, 11, 20, 22, 44, 55, 110 whose sum is 284, while those of the second are 1, 2, 4, 71, 142 whose sum is 220.

No general theorems concerning the forms of amicable numbers, similar to those which we have developed for perfect numbers, have thus far been proved. A number of pairs of amicable numbers have been found by various people; but the subject is still in a very rudimentary stage. It does not seem to have much mathematical importance, but rather to be one of the curiosities of the theory of numbers.

There are a good many other curiosities in this field of mathematics which have aroused interest in the course of time, and some of which have proved afterwards to be of significance. Let us look at two of these.

(a) *Farey's Series.* A brief history and a short account of these "series"[2] is found in the address by G. H. Hardy which was referred to at an earlier point (see p. 175). They are built up by writing down all rational numbers which do not exceed 1 and whose

[1] Compare Dickson, *Introduction to the Theory of Numbers*, p. 5.

[2] The word "series" is here to be understood as "finite sequence" (compare footnote on p. 58), contrary to accepted usage.

denominator does not exceed a preassigned *integer* k, in the order of their magnitude, each reduced to lowest terms, and excluding duplicates. For instance, if we take $k = 5$, we have to arrange in order of magnitude the rational numbers $0, 1, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}$. Thus we obtain the finite sequence $\frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1}$. In a similar way, we can write down a Farey Series for any other preassigned integer k.

Suppose now that $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$ are two consecutive terms in any such "series." Then it will always be true that $a_2b_1 - a_1b_2 = 1$. The reader will readily verify this property for the special case $k = 5$ written down above. A proof that it is always true can be found in Hardy's paper; we shall not reproduce it here.

An interesting and immediate consequence of this remarkable fact is the following property of the Farey Series: Suppose that $\frac{a_1}{b_1}, \frac{a_2}{b_2}$ and $\frac{a_3}{b_3}$ are three consecutive terms of a Farey Series. Then $a_2b_1 - a_1b_2 = 1$ and $a_3b_2 - a_2b_3 = 1$, so that $a_2b_1 + a_2b_3 = a_1b_2 + a_3b_2$. From this we conclude that $a_2(b_1 + b_3) = b_2(a_1 + a_3)$ and hence that $\frac{a_2}{b_2} = \frac{a_1 + a_3}{b_1 + b_3}$, i.e.: *Of three consecutive terms in a Farey Series the middle term is equal to the fraction obtained by adding together the numerators and the denominators of the other two terms.* The reader will certainly want to verify this fact for a few special Farey Series, like the one for $k = 5$, which was written down above.

The Farey Series which were for a long time little more than an interesting curiosity have in recent years been shown to be of importance in connection with some of the fundamental problems in the theory of numbers. The same can not be said of the following topic; it is still very largely a subject for recreation and for that reason a fitting one with which to end a chapter which promised some amusement.

8	1	6
3	5	7
4	9	2

FIG. 28

(b) *Magic Squares.* The games of hopscotch, played on the city streets, and of shuffleboard, played on ocean liners, have a great many addicts. It is probably known to everyone that the central part of the court used in these games consists of a square in which are marked off the integers from 1 to 9, as in Fig. 28.

But perhaps not every player of these games has observed that the sum of the three numbers in *every* horizontal line, and in *every* vertical line is equal to 15.

Any arrangement of the natural numbers from 1 to n^2 in a square, in such a way that the sum of the n numbers in every horizontal line, and in every vertical line, is equal to that in every other, is called a *magic square of order n*.[1]

Other examples of magic squares, besides the one exhibited in Fig. 28, are:

4	8	3
9	1	5
2	6	7

15	10	3	6
4	5	16	9
14	16	2	7
1	8	13	12

4	7	15	18	21
6	14	17	25	3
13	16	24	2	10
20	23	1	9	12
22	5	8	11	19

FIG. 29

in which the constant sums are equal respectively to 15, 34 and 65.[1]

There is a quite interesting literature attached to this subject of magic squares, dealing with their use in various rituals, the power attributed to them, the methods of constructing them, etc. The reader who would like to follow this subject further will find material and references in vol. 17 of the *Encyclopaedia Britannica*, 11th edition; in Smith and Mikami, *The History of Japanese Mathematics*, page 117, and in W. S. Andrews, *Magic Squares and Cubes*.[2] The latter book reproduces an engraving by Albrecht

[1] If we observe that the numbers from 1 to and including n^2 form an arithmetic progression, and if we recall the fact referred to in the footnote on page 24, we will recognize that the sum of all the numbers in a magic square of order n is equal to $\frac{n^2(n^2+1)}{2}$. Consequently the sum of the numbers in any one horizontal or vertical line is equal to $\frac{n(n^2+1)}{2}$. For $n = 3$, this number reduces to 15, which is indeed the number found for the magic square of Fig. 28 and for the first in Fig. 29; for $n = 4$, we have $\frac{n(n^2+1)}{2} = \frac{4 \cdot 17}{2} = 34$, and for $n = 5$, this expression becomes $\frac{5 \cdot 26}{2} = 65$.

[2] See also the footnote on p. 222.

Dürer called "Melancholy," in the background of which appears the square

16	3	2	13
5	10	11	8
9	6	7	12
4	15	14	1

FIG. 30

A few years ago a paper appeared which connected a certain method for constructing magic squares with some of the concepts of the theory of numbers with which we have become acquainted.[1] We shall give an outline of a small portion of the contents of this paper.

Let us consider a square consisting of n^2 cells. The first problem consists in finding a systematic way of filling up these cells with the natural numbers from 1 to and including n^2. For this purpose we use the "uniform step" method. Let us designate each cell by its two "coördinates," counted from the lower left-hand corner; thus (p, q) designates the compartment in the pth column and the qth row. A step $[a, b]$ consists then in going from (p, q) to $(p+a, q+b)$. If one or both of the numbers $p + a$ and $q + b$ should exceed n, we replace them by their residues, mod. n, in order that we may remain in the given square. We agree now upon some fixed step $[a, b]$ and start by putting 1 in an arbitrary cell (p, q); then we place 2, 3, . . . i.e. in the successive cells obtained by starting from (p, q) and taking the step $[a, b]$ in going from any cell to the next. Suppose, for example, that $n = 7$, that we agree on the step $[2, 3]$ and that we put 1 in the cell $(3, 4)$, compare Fig. 31. Then 2 should go in the cell $(5, 7)$, 3 in $(7, 10)$ i.e. in $(7, 3)$, 4 in $(9, 6)$ i.e. $(2, 6)$, etc. In general, the number k will occupy the cell

$$(p + (k - 1)a, q + (k - 1)b);$$

and, in particular, the number $n + 1$ will be assigned to the cell $(p + na, q + nb)$. But, since $p + na \equiv p$, mod. n and $q + nb \equiv$

[1] Compare D. N. Lehmer, "On the Congruences Connected With Certain Magic Squares," *Transactions of the American Mathematical Society*, Vol. 31, 1929, p. 529.

		11		2		
	4					13
			8(15)		6	
10		1 (8)				
				12		3
	14		5			
7					9	

FIG. 31

q, mod. n, this would be the cell (p, q), which is already occupied by 1. In the illustration of Fig. 31, in which $n = 7$, we see that indeed 8 claims the same cell as 1. This difficulty has to be avoided. But before doing so, let us inquire whether such a conflict can occur before we reach the number $n + 1$. In other words, if

$$1 \leqq k_1 < n + 1 \quad \text{and} \quad 1 \leqq k_2 < n + 1,$$

will it ever happen that

$$p + (k_1 - 1)a \equiv p + (k_2 - 1)a$$

and

$$q + (k_1 - 1)b \equiv q + (k_2 - 1)b, \text{ mod. } n?$$

Clearly, if it did, we would have

$$(k_1 - k_2)a \equiv 0, \text{ mod. } n, \quad \textit{and} \quad (k_1 - k_2)b \equiv 0, \text{ mod. } n.$$

Now, if a and b are relatively prime to n, this can only happen if $k_1 - k_2 \equiv 0$, mod. n; but, if k_1 and k_2 are at most equal to n and at least equal to 1, this can only occur if $k_1 = k_2$. Therefore if a and b are relatively prime to n, there will be cells available for all the numbers from 1 up to and including n.

But what shall we do with $n + 1$? We can not put it in the cell already occupied by 1; neither can it be placed in the cell obtained from the one chosen for 1 by the uniform step $[a, b]$ used thus far.

We resolve the difficulty by introducing a new step (the reader will observe the resemblance of this subject with another favorite amusement), called the break-step. Let us denote it by $[a_1, b_1]$. Then we shall place $n + 1$ in the cell obtained from the cell (p, q) occupied by 1, by using the break-step $[a_1, b_1]$, that is to say in the cell $(p + a_1, q + b_1)$. Obviously if this move is to avoid conflicts effectively, we must not only choose a_1 and b_1 in such a way that $[a_1, b_1]$ is a step different from $[a, b]$, but also so that it is different from several steps $[a, b]$ made successively. This means that $\frac{a_1}{a}$ must be different from $\frac{b_1}{b}$, or that $a_1b - ab_1 \neq 0$. Having chosen the break-step in this way, and having placed $n + 1$ in the cell $(p + a_1, q + b_1)$ we proceed again with the uniform step $[a, b]$ to locate $n + 2, n + 3, \ldots 2n$. For our special example, let us take $a_1 = b_1 = 1$. The numbers 8, 9, 10, 11, 12, 13, 14 will then occupy the positions indicated in Fig. 31.

The next number to cause trouble is $2n + 1$, because it would come in the same cell as $n + 1$; to avoid unpleasantness, we put $2n + 1$ in the cell reached from that occupied by $n + 1$ by using the break-step $[a_1, b_1]$ once more and then go forward again, with the uniform step $[a, b]$. In this way, using the uniform step $[a, b]$ whenever we can do so without difficulty and the break-step $[a_1, b_1]$ to remove conflicts, we continue until all the cells have been used up. Will we indeed escape conflicts in this way, and will the result be a magic square? The answers to these questions are given in the paper by Lehmer, which was mentioned on page 219, viz.:

If the numbers a, b, a_1, b_1, and $ab_1 - a_1b$ are all prime to n, then the arrangement of the numbers from 1 to and including n^2 obtained by starting at an arbitrary cell (p, q), and by using the uniform step $[a, b]$ and the break-step $[a_1, b_1]$ in the manner described above, will always be a magic square of order n.

The complete proof of this statement would carry us a little too far afield; the reader who has followed the discussion should be able to appreciate the reasonableness of the conditions which the conclusion specifies. To justify the general statement, he has to carry through the successive steps in the reasoning. Not every reader will be sufficiently interested in the subject to undertake this task; no one will have any difficulty in applying the method to special cases.

For the example started in Fig. 31, the final result is as follows:

20	40	11	31	2	22	49
33	4	24	44	15	42	13
46	17	37	8	35	6	26
10	30	1	28	48	19	39
23	43	21	41	12	32	3
36	14	34	5	25	45	16
7	27	47	18	38	9	29

The reader will find it a simple matter to verify that the sum of the numbers in any horizontal line, or in any vertical line, is always equal to 175 (compare footnote 1 on p. 218).

Thus we have become acquainted with one method for constructing magic squares. Is it effective in every case, that is to say, can the conditions which it requires always be fulfilled? What other schemes can be used? Are there more mysteries hidden in these arrangements of numbers, and are there other such inventions? If the reader wishes to find answers to such questions, he can begin by entrusting himself to one of the guides mentioned on page 218.[1]

But our wanderings through the theory of numbers must now come to a close. For it is high time that we turn to some of the other things in the vast domain of mathematics which are within our range. New prospects are to be opened up in the following chapters. Before we enter upon these new adventures, we had better consider the positions to whose conquest we have devoted the last pages. This is the purpose of the next section.

[1] There are a good many books which deal with the subjects discussed in this chapter. To those which have already been mentioned, we add the following: H. E. Dudeney, *Modern Puzzles;* W. W. R. Ball, *Mathematical Recreations and Problems;* C. G. Bachet (compare p. 166), *Problèmes plaisants et délectables;* W. Ahrens, *Hebräische Amulette mit magischen Zahlenquadraten;* W. Ahrens, *Mathematische Unterhaltungen und Spiele,* 2 vols.; W. Lietzmann, *Lustiges und merkwürdiges von Zahlen und Formen.* Of a different, although somewhat related character are the following two books, which contain a large number of anecdotes and quotations relating to mathematics: R. E. Moritz, *Memorabilia Mathematica;* W. Ahrens, *Scherz und Ernst in der Mathematik.* We mention finally a book which connects not only with the subject of this chapter but also with topics which occur in earlier chapters, viz. H. A. Merrill, *Mathematical Excursions.*

103. A curiosity shop.

1. Prove that if $2^k + 1$ is a prime number, then the sum of the divisors of the number $2^{k-1}(2^k + 1)$ excluding this number itself is 2 less than the number itself. Verify this result for special cases.

2. Show that 12, 18, 20, 28 and 32 are the first five numbers in order of magnitude which have exactly 6 divisors.

3. Prove that if the number of divisors of a number N is equal to a prime number p, then N has the form a^{p-1}, where a is also a prime.

4. Determine the form which a number N must have if the number of its divisors is equal to the product of 2 primes.

5. Determine the possible forms of a number N which has exactly 24 divisors; which is the least of these numbers?

6. Show that if the number of divisors of N is equal to n, then the product of these divisors is equal to $N^{\frac{n}{2}}$.

7. Prove that a number N is a perfect square if and only if the number of its divisors is odd.

8. Determine the form of a number whose square is equal to the product of its divisors.

9. Determine the form of a number whose cube is equal to the product of its divisors.

10. Show that an even perfect number must end in 6 or 8.

11. Show that the sequence of residues, mod. 9, of the successive powers of 2, beginning with 2^1, consists of the numbers 2, 4, 8, 7, 5, 1, repeated indefinitely.

12. Show that if $2^k - 1$ is a prime number > 3, its residue, mod. 9, is 7, 4 or 1.

13. Use the results of 11 and 12 to show that the residue, mod. 9, of an even perfect number, which exceeds 6, must be 1.

14. Write the first four even perfect numbers first in the decimal scale, then in the binary scale.

15. Determine the only possible general form of an even perfect number in the binary scale.

16. Derive the condition which must hold if the sum of all the divisors of a number N is equal to three times that number.

17. Show that if $N = p^k$, where p is a prime then N can not satisfy the condition of 16.

18. Show that the sequence of numbers obtained from a Farey Series by increasing all the terms by an integer n has also the fundamental property of the Farey Series.

19. Show that the sequence of numbers obtained from the Farey Series for k by increasing all the terms by a rational number $\frac{p}{q}$ has the

property, that if $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$ are two consecutive terms of this sequence then $a_2b_1 - a_1b_2 = q^2$, provided q is relatively prime to all the integers less than or equal to k (in particular if q is a prime greater than k); show also that if $\frac{a_1}{b_1}, \frac{a_2}{b_2}, \frac{a_3}{b_3}$ are three consecutive terms of such a sequence, then $\frac{a_2}{b_2} = \frac{a_1 + a_3}{b_1 + b_3}$.

20. Prove that two adjacent terms in the Farey Series for $k > 1$ can not have the same denominators.

21. Prove that if two adjacent terms in a Farey Series have the same numerator, that numerator must be 1 and the corresponding denominators must differ by 1.

22. Prove that if $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$ are consecutive terms in a Farey Series and $b_2 > b_1$, then $a_2 > a_1$, while if $b_2 < b_1$, then $a_2 \leqq a_1$. (In virtue of 21, the equality sign can occur only if $a_2 = a_1 = 1$.

23. Show that if $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$ are any two non-consecutive terms in the Farey Series for k, such that $\frac{a_1}{b_1} < \frac{a_2}{b_2}$, then $\frac{a_1 + a_2}{b_1 + b_2}$ also belongs to that sequence and lies between the two given terms, provided $b_1 + b_2 \leqq k$.

24. Prove that if $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$ are two consecutive terms of the Farey Series for k then $b_1 + b_2 > k$.

25. Prove that the Farey Series for $k + 1$ can be obtained from that for k by inserting the term $\frac{a_1 + a_2}{b_1 + b_2}$, between any two successive terms $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$ for which $b_1 + b_2 = k + 1$.

26. Show that if p_b denotes the number of numbers which do not exceed b and are prime to b, then the number of terms in the Farey Series for k is equal to $2 + p_2 + p_3 + \cdot \cdot \cdot + p_k$.

27. Show that a magic square remains magic if any two horizontal lines, or any two vertical lines, are interchanged; also if the horizontal lines are written vertically and the vertical lines horizontally.

28. Show that all the magic squares of order 3 which can be constructed by means of the uniform step and break-step method are obtainable from any one of them by the interchanges mentioned in 27.

29. Show that no magic square of order 4 can be constructed by the method of uniform step and break-step; generalize this statement so as to include all squares of even order.

30. Show that two magic squares which are constructed with the same uniform step and the same break-step, and which differ only in the position chosen for 1, can be obtained from each other by the changes described in 27.

31. Construct at least two magic squares of order 5 which can not be converted into each other by means of the interchanges of 27.

32. Is it possible to construct a magic square of order 5 such that the nine inner cells form a magic square of order 3?

CHAPTER X

A GARDEN OUTSIDE THE WALL

Geometry became one of the most powerful expressions of that sovereignty of the intellect that inspired the thought of those times. At a later epoch, when the intellectual despotism of the church, which had been maintained through the middle ages, had crumbled, and a wave of scepticism threatened to sweep away all that had seemed most fixed, those who believed in truth clung to Geometry as to a rock, and it was the highest ideal of every scientist to carry on his science "more geometrico." — H. Weyl, *Space, Time and Matter*, p. 1.

104. Memories and prospects. The part of the journey that will occupy us now has for its purpose to renew acquaintance with a number of things with which we have been familiar since school-days, in order to get some idea of the extensive domain of which these are but a part.

We connect with the subject indicated on school programs as plane geometry. This, in the experience of most people, consists of a number of theorems and proofs to be learned, of facts to be accepted, of definitions to be memorized, of exercises to struggle with. For some perhaps, it has opened up avenues of wide prospect, opportunities for the exercise of inventive thought. It is in such parts that we shall want to travel, for thus may we discover new scenes of beauty and of interest, and perhaps gain some insight into the principles which operate in their development.

It is as if we were returning after many years' absence to a garden which we have known in early childhood, and which was then limited by walls too high for us to climb. Coming back to it as men and women, the walls no longer obstruct our view, the garden stretches far beyond what we remember from our early experiences, apparently to an unlimited extent. Its lanes and avenues appear to be laid out according to some plan, not grasped very clearly at first, but yielding more and more to our attempts at understanding. New plants, rare flowers are seen on every hand, and as we walk on we wonder what determines their growth, whence comes the energy which nourishes these wonderful creations, which are the seeds from which they spring. And we ask why we were not aware

of all this when we were children. It seems unbelievable that it is not so much the garden which has changed, but more that we have grown up, that these marvelous things were there then, but beyond the reach of our eyes, beyond the range of our wanderings.

Let us recall a few familiar facts.

105. Lines and circles. When four points A, P, B and Q are taken on the circumference of a circle so that when we go around the circle either in one sense, or else in the other, these points will occur in the order named, we say that the angle APB intercepts the arc AQB of the circle (see Fig. 32), and the angle AQB intercepts the arc APB. Among the properties of this figure which we learned in school are the following:

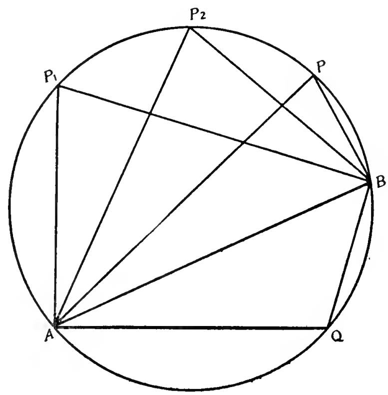


FIG. 32

(a) The sum of the angles APB and AQB is equal to a straight angle, i.e. an angle of 180°, and, conversely, if P and Q lie on opposite sides of the line AB and $\angle APB + \angle AQB = 180°$, then the circumference determined by A, B and P also passes through Q.

(b) If the angles AP_1B and AP_2B intercept the same arc on a circle, then they are equal; and, conversely, if P_1 and P_2 lie on the same side of the line AB, and $\angle AP_1B = \angle AP_2B$, then the circumference determined by A, B and P_1 also passes through P_2.

(c) An angle which intercepts a semi-circle is equal to a right angle, i.e. an angle of 90°.

(d) There always exists one and only one circle which passes through three distinct points A, P, B, arbitrarily placed, provided they do not lie on one straight line; if $\angle APB$ is a right angle, then AB is a diameter of this circle.

There are many other things which will have to be recalled as we go on; it is inconvenient and unnecessary to make a list of them. The reader will readily have access to a book on plane geometry which will enable him to refresh his memory.

We shall, therefore, pass on at once to a simple and interesting

fact which follows quite easily from the things we have just mentioned.

Theorem L. The feet of the three perpendiculars dropped from a point P on the circumference of a circle upon the sides BC, CA and AB of a triangle whose vertices A, B and C lie on this circumference, lie on a straight line (see Fig. 33).

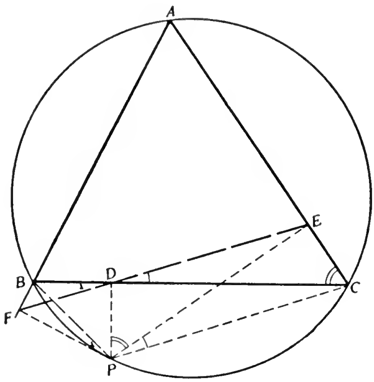


FIG. 33

Proof. The feet of the perpendiculars are denoted by D, E and F. Since PDC is a right angle, the circle determined by P, D and C has PC as a diameter; similarly the circle determined by P, E and C has PC as a diameter. But there is only one circle which has PC as a diameter; consequently P, D, E and C lie on one circumference. On this circle, which is not drawn in the figure, the angles EDC and EPC intercept the same arc; and the angles ECD and EPD also intercept the same arc. Therefore it follows, as indicated in the figure, that

$$(10.1) \qquad \angle EDC = \angle EPC \quad \text{and} \quad \angle ECD = \angle EPD.$$

The same reasoning, applied to the points P, D, B and F, will show that, as marked in the figure,

(10.11)

$$\angle BDF = \angle BPF; \text{ moreover } \angle DPF = 180° - \angle DBF = \angle ABC.$$

Finally, since A, B, P and C also lie on a circumference,

$$(10.12) \quad \angle BPC = 180° - \angle BAC = \angle ABC + \angle ACB.$$

A little simple algebra will now lead us to the desired result. For, from (10.12) and the second part of (10.1), we conclude that

$$\begin{aligned}\angle BPD + \angle EPC &= \angle BPC - \angle EPD \\ &= \angle ABC + \angle ACB - \angle ECD = \angle ABC;\end{aligned}$$

and the second part of (10.11) is equivalent to the statement

$$\angle BPD + \angle BPF = \angle ABC.$$

Consequently $\angle EPC = \angle BPF$, and therefore, by use of the first parts of (10.1) and (10.11)

$$\angle EDC = \angle BDF.$$

This carries with it the conclusion that FDE is a straight line, as is asserted in the theorem.

There are many other facts connected with this theorem; it can be made the starting point of a considerable expedition. But it is enough for our purpose to have opened a vista upon such further ventures. The reader will find plenty of information concerning the line which this theorem has shown to exist in books that are devoted to geometry. The line is usually called the *Simson* line of P with respect to the triangle ABC,[1] apparently erroneously attributing its discovery to the English geometer Robert Simson (1687–1768).

We pass on to another remarkable result which lies within easy reach. While as has already been recalled it is the rule that *three* points, which do not lie on one line, do lie on a circle, it is something worth noticing whenever *more* than three points lie on one circumference. This lends particular interest to the next theorem. Before stating it, we had better make sure that we know the meaning of the technical terms that occur in its statement and the additional propositions needed in its proof.

A perpendicular from a vertex of a triangle to the opposite side is called an *altitude* of the triangle. Every triangle has therefore three altitudes. They pass through one point; that point is called the *orthocenter* of the triangle. Also, the line joining the midpoints of any two sides of a triangle is parallel to the third side; and if two lines are parallel to the two sides of a right angle, they are perpendicular to each other. Now we can proceed.

Theorem LI. The three midpoints of the sides of a triangle, the feet of the three altitudes, and the three midpoints of the segments from the orthocenter to the vertices, these *nine* points lie on one circumference.

Proof. Figure 34 indicates the positions of the points with which we are concerned: D, E and F are the midpoints of the sides, L, M and N are the feet of the altitudes; P, Q and R are the midpoints of the segments from the orthocenter, H, to the vertices

[1] See R. A. Johnson, *Modern Geometry*, p. 137; N. Altshiller-Court, *College Geometry*, p. 115.

A, B and C. We have to prove that the nine points D, E, F; P, Q, R; L, M, N lie on one circumference.

(a) Since P and E are the midpoints of the sides AH and AC of $\triangle AHC$, we conclude that $PE \parallel HC$; and since E and D are the midpoints of the sides AC and BC of $\triangle ABC$, we conclude that $ED \parallel AB$. Therefore PE and ED are parallel to the two sides of the right angle CNA; consequently they are perpendicular to each other, i.e $\angle PED$ is a right angle. This being so, the circle determined by P, E and D has PD as a diameter.

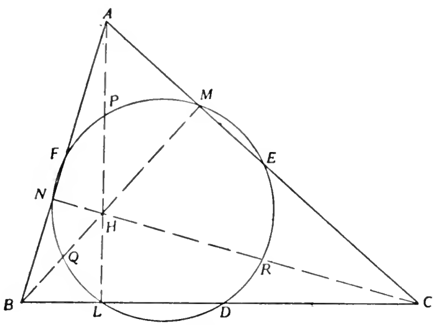


FIG. 34

We repeat this argument with the points P, F and D: $PF \parallel BH$, $FD \parallel AC$, the lines PF and FD are then parallel to the sides of the right angle BMC, therefore perpendicular; consequently the circle determined by P, F and D has PD as diameter.

On the other hand, it is obvious that the circle determined by P, L and D also has PD as a diameter.

But since there is only one circle which has PD as a diameter, we conclude that the points P, L, D, E and F all lie on one circumference.

(b) If we start all over again, using now the points Q and E in place of P and D, we arrive by the same argument at the conclusion that the points Q, M, D, E and F lie on one circumference. Finally, going through the same reasoning once more starting from the points R and F, we reach the conclusion R, N, D, E and F lie on one circumference. These three circumferences all pass through the three points D, E and F. But there is only one circumference through these three points, for they certainly do not lie on a straight line. Therefore the three circumferences which we have obtained can not be different from each other; they are one and the same thing, parading under different names. *Ergo*, the nine points P, Q, R; L, M, N; D, E, F lie on one circumference. It has the lines PD, QE and RF as diam-

eters; therefore these three lines meet in a point, viz. the center of the circumference.

Not only have we proved our theorem, we have also learned something about the location of the center of the circle of which it asserts the existence. It bears the obvious name of *nine point circle*. Numerous other results lie close at hand; the books mentioned on page 229 will supply the interested reader with many further details.

The two theorems which have been presented may give the reader some idea of the sort of thing to be found in geometry if we pursue it far enough. There is a great variety of curious and interesting facts already known; many more, apparently in endless succession, will appear if inventive power is applied to the study of established results. Even if we limit ourselves to the configurations familiar from elementary plane geometry, which consist of straight lines and circles only, there is a limitless field for study and investigation. But judgment and insight are required if the results attained are to form more than a mere agglomeration of isolated facts, if they are to be organically related. In order to give some idea of such unity, we will develop in the next few sections a number of theorems which are interesting in themselves and interrelated in a significant way.

106. Jacobi guides once more. Very useful as a guide in the development of a mathematical theory, is a second principle usually attributed to Jacobi (compare p. 40), viz.: "Man muss immer generalisieren"—we must always generalize. An example is ready at hand in connection with Theorem L.

The lines PD, PE and PF in Fig. 33 are perpendicular to the sides BC, CA and AB of $\triangle ABC$ respectively; therefore the six angles which these lines make with the sides of the triangle are all equal to each other. We obtain a more general situation if we draw through P three lines, which make three equal angles with the sides of the triangle; the three remaining angles will then also be equal, but not necessarily equal to the first three. And it turns out to be possible to prove that also in this case we obtain three collinear points.

Theorem LII. If through a point P on the circle determined by three points A, B and C, lines PL, PM and PN be drawn such that the angles PLC, PMC and PNB are equal, then the points L, M and N lie on a straight line (see Fig. 35).

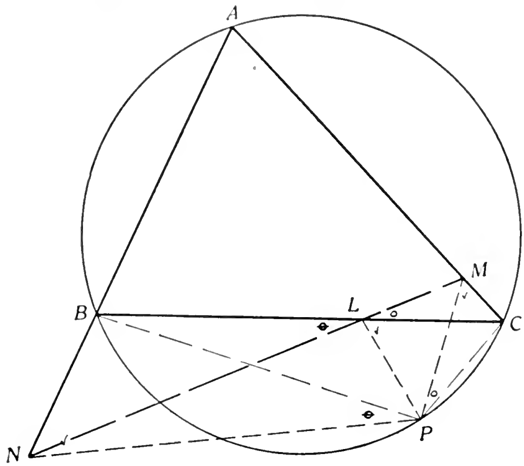


FIG. 35

Proof. The proof is almost identical with that of Theorem L; we have to make use, however, of the converses stated in (a) and (b) of *105*. From them it follows:

1. that the points P, C, M and L lie on one circumference, so that $\angle MLC = \angle MPC$, and $\angle LPM = \angle LCM$;

2. that the points P, L, B and N lie on one circumference so that

$$\angle BLN = \angle BPN,$$

and $$\angle BPN + \angle BPL = 180° - \angle NBL = \angle ABC.$$

Moreover,

$$\angle BPC = 180° - \angle BAC = \angle ABC + \angle ACB,$$

so that, using the second part of 1:

$$\begin{aligned} \angle BPL + \angle MPC &= \angle BPC - \angle MPL \\ &= \angle ABC + \angle ACB - \angle MPL \\ &= \angle ABC + \angle ACB - \angle MCL = \angle ABC; \end{aligned}$$

consequently, by means of the second part of 2,

$$\angle BPL + \angle MPC = \angle BPN + \angle BPL,$$

so that $$\angle MPC = \angle BPN;$$

from this, we conclude on the strength of the first parts of 1 and 2, that $\angle MLC = \angle BLN$, so that the points M, L and N lie on a straight line.

Thus we see that Theorem L is but a special case of a more general theorem. While it is, of course, of interest to know that the more general theorem is valid, from a certain point of view it is more important to *ask* whether it may not be true than actually to answer the question. To a considerable extent scientific progress depends upon the asking of fruitful questions, that is to say upon intellectual curiosity. Many further questions can be asked relative to Theorem L, the following, for example: Will the theorem still be true if instead of a triangle inscribed in a circle we consider a polygon? Will it be true if instead of a circle there be introduced another closed curve, an ellipse, for example? If the answer to either of these questions should be in the negative, we would try to account for such a deviation from the previous conclusion.

Remark. Attention should be called to a difficulty which arises in connection with Theorem LII. The lines PL, PM and PN make two angles with the sides BC, CA and AB of the triangle ABC respectively; and it has not been made clear which angle of each pair is to be used. The point is obviously an important one. We shall not deal with it however, because it is too technical in character; it is discussed in books that go in for a thorough treatment of these problems.[1]

Another aspect of the method of generalization and unification will be brought out when we approach Theorem LII from a different side. We begin with an apparently unrelated question.

107. A connecting link.

Theorem LIII. If three points L, M and N be taken on the sides of a triangle ABC, then the three circles determined by A, M, N, by B, N, L and by C, L, M have one point in common (see Fig. 36).

Proof. We take the points L, M and N on the segments BC, CA and AB, thus leaving out of consideration the possibility that one or more of them might lie on the extensions of these segments. The circles determined by A, M, N and by B, N, L will then, since they have the point N in common, have a second point P in common. It then follows that $\angle NPM = 180° - A$; and that $\angle NPL = B$. Therefore $\angle MPL = \angle NPM - \angle NPL = 180° - A - B = C$.

[1] See, e.g., Johnson, *Modern Geometry*, pp. 11–15.

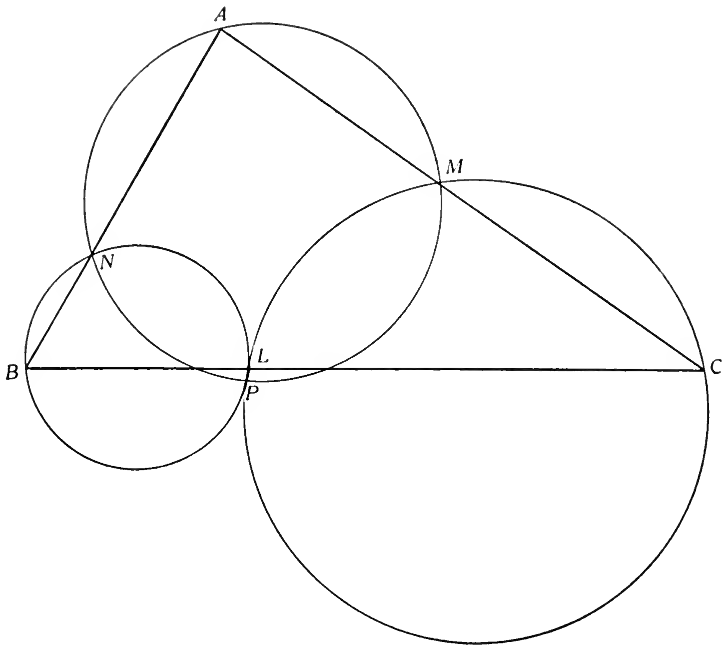


FIG. 36

But this implies in virtue of (a), *105*, that P lies on the circumference determined by C, L, M. The point P lies then on each of the three circumferences, so that our proposition is proved, at least for the case here considered.[1]

The cases arising when one or more of the points lie on the extensions of the sides of the triangle can be treated in an entirely similar manner and will be useful exercises for the reader (see *108*, 13, 14).

This theorem is usually attributed to A. Miquel (first half of 19th cent.).[2] The point P is called the *Miquel point* for the set of points L, M, N with respect to the triangle ABC.

By special choice of the positions of L, M, N a number of par-

[1] If the point P falls inside the triangle determined by L, M and N, the argument, in this proof and in the corollaries which follow, has to be slightly modified; the conclusions remain unchanged.

[2] But see Johnson, *op. cit.*, p. 133.

ticular cases of the theorem come within range. We are not going to aim at these; it is our purpose to make a connection with Theorem LII.

From a consideration of Fig. 36, we obtain the following additional results:

$$\angle PLC = \angle PMC,$$

and $$\angle PLC = 180^\circ - \angle PLB = \angle PNA.$$

$$\therefore \angle PLC = \angle PMC = \angle PNA.$$

Corollary 1. If P is the Miquel point for the set of points L, M, N with respect to the $\triangle ABC$, then the angles PLC, PMC and PNA are equal.

If, conversely,

$$\angle PLC = \angle PMC = \angle PNA,$$

then P, L, M, C lie on one circumference; and $\angle PLB = \angle PNB$, so that P, L, N, B lie on one circumference; hence P is the Miquel point of the set of points L, M, N with respect to $\triangle ABC$. This gives us the converse of Cor. 1, viz.:

Corollary 2. If the lines PL, PM and PN are such that $\angle PLC = \angle PMC = \angle PNA$, then P is the Miquel point of the set of points L, M, N with respect to $\triangle ABC$.

Next we observe that

$$\angle NML = \angle NMP - \angle PML = \angle BAP - \angle PCB,$$

and that the latter angles depend only upon the position of the point P. Similarly

$$\angle MLN = 360^\circ - \angle MLP - \angle PLN = \angle ACP + \angle PBA$$

and $$\angle LNM = \angle MNP - \angle PNL = \angle CAP - \angle PBC.$$

Consequently the angles NML, MLN and LMN are entirely determined by the position of the point P. We draw the following conclusion:

Corollary 3. If two sets of points L, M, N and L_1, M_1, N_1 have the same Miquel point with respect to a triangle ABC, then the triangles LMN and $L_1M_1N_1$ have their corresponding angles equal and are therefore similar.

Starting from an arbitrary point P we can construct various sets of three points L, M, N each, such that P is the Miquel point with

respect to $\triangle ABC$ for every one of these sets. We have but to construct an arbitrary circle through P and A, cutting the sides AB and AC in N and M respectively, and then construct the circle determined by P, M and C, or the circle determined by P, N and

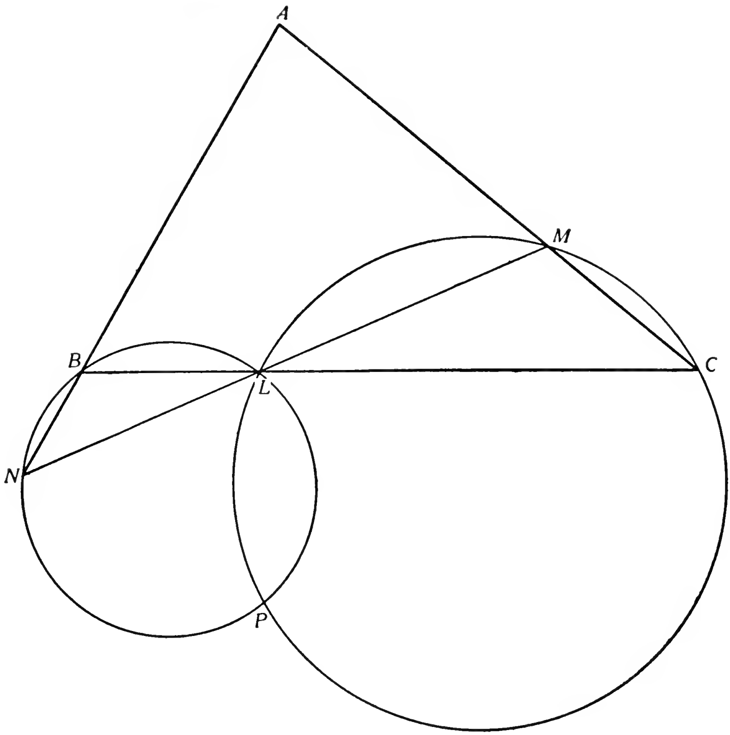


FIG. 37

B; both these circles will intersect the side BC in the third point L of the set.

The desired connection between the Miquel theorem and Theorem LII can now be established by the following theorem.

Theorem LIV. If three points L, M, N on the sides BC, CA and AB of $\triangle ABC$ lie on a straight line, then the corresponding Miquel point lies on the circle determined by A, B and C; conversely, any three points L, M, N, for which the Miquel point with respect to a

$\triangle ABC$ lies on the circle determined by A, B and C, lie on a straight line.

Proof. In the proof of this theorem we have to use cases of the Miquel theorem (Theorem LIII) which were left for the reader's private delectation. The importance of a complete proof of this theorem thus becomes evident for a satisfactory discussion of the question before us. We proceed then as follows:

(a) Suppose that L, M, N lie on a straight line (see Fig. 37) and consider the triangle AMN. On the side AM, we choose the point C; on the side MN the point L, and on the side NA the point B. Then we know, from the Miquel theorem, as applied to $\triangle AMN$ and the set of points C, L, B that the three circumferences determined by A, B, C, by M, C, L and by N, L, B meet in a point. But the point of intersection of the last two circumferences is the Miquel point of L, M, N with respect to $\triangle ABC$; therefore this point lies on the circle determined by A, B and C.

(b) Suppose now that P is on the circle determined by A, B and C, and that L, M, N are three points for which P is the Miquel point with respect to $\triangle ABC$.

We construct then first a special set of three points, L_1, M_1, N_1 for which P is the Miquel point with respect to $\triangle ABC$ (compare the paragraph following Cor. 3 above), by drawing the circle on PC as a diameter, cutting AC and BC in M_1 and L_1 respectively, and then constructing the circle determined by P, L_1 and B cutting AB in N_1. Since PC is the diameter of the first of these circles $\angle PM_1C = 90°$ and $\angle PL_1C = 90°$; therefore $\angle PL_1B$ is also a right angle,

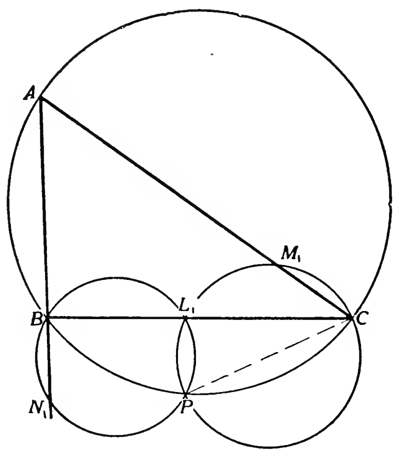


FIG. 38

BP is then the diameter of the second circle and consequently $\angle PN_1B = 90°$. The points L_1, M_1 and N_1 are then the feet of the perpendiculars upon the sides of $\triangle ABC$ from a point in its circumscribed circle; we know from Theorem L that L_1, M_1

and N_1 lie on a straight line, viz. the Simson line of P. If L, M, N is now an arbitrary set of three points for which P is the Miquel point, then we know from Cor. 3 that the triangles LMN and $L_1M_1N_1$ must have their corresponding angles equal. Hence L, M, N must also lie on a straight line. This completes the proof of Theorem LIV. The road to Theorem LII should now be clear. For if PL, PM, PN make equal angles with the sides of $\triangle ABC$, P is the Miquel point of L, M, N (Cor. 2 on p. 235); if moreover this point P is on the circumference determined by A, B and C, the points L, M, N are on a straight line. This was the assertion of Theorem LII.

Thus, starting from an apparently unrelated source, we have reached a familiar result; the whole group of theorems and corollaries, from L to LIV has been unified. These processes of generalization and unification are fundamental for the whole field not only of geometry, but of all mathematics. It is by applying them on an ever extending scale that insight and understanding are gained. In the following set of exercises the reader has an opportunity to make himself more familiar with the theorems we have been discussing.

108. Laying out more paths.

1. Prove that if A and P are the extremities of a diameter of a circle, B and C arbitrary points on its circumference, then BC is the Simson line of P with respect to $\triangle ABC$.

2. Show that the diameter of the nine-point circle of a triangle ABC is equal to the radius of the circle determined by A, B and C.

3. Prove that the four circumferences determined by the vertices of the four triangles formed by omitting any one of four lines which intersect two by two and no three of which pass through the same point, have one point in common.

4. Prove that the Miquel point which corresponds to the feet of the altitudes of a triangle is the orthocenter of the triangle.

5. Prove that if P and Q are the extremities of a diameter of a circle, and A, B, C arbitrary points on its circumference, then the Simson lines of P and Q with respect to $\triangle ABC$ are mutually perpendicular.

6. Prove that the centers of the circles ANM, BLN and CML constructed in Fig. 36, are the vertices of a triangle whose angles are equal to those of $\triangle ABC$. (Remember that the line joining the centers of two circles is perpendicular to their common chord.)

7. Prove that the perpendiculars PL, PM and PN from a point P, on the circle determined by the points A, B, C, to the sides BC, CA and

AB of the $\triangle ABC$, meet the circle again in points A_1, B_1 and C_1 respectively, such that the triangles $A_1B_1C_1$ and ABC have equal angles.

8. Prove that if the lines AP and BP meet the sides BC and CA of $\triangle ABC$ respectively in L and M, and the points P, M, C and L lie on one circumference, then the two circumferences determined by A, P, M and by B, P, L have a point of AB in common.

9. Prove that, in Fig. 33, the triangles PDE and PBA have equal angles; indicate two other similar pairs of equiangular triangles.

10. Prove that if the perpendiculars from an arbitrary point P meet the sides of a triangle ABC in points L, M and N then P is the Miquel point of L, M, N with respect to the triangle.

Note. The triangle formed by the points L, M and N is called the *pedal triangle* of P with respect to $\triangle ABC$. (What becomes of the pedal triangle when P lies on the circumference determined by A, B and C?)

11. Prove that if the lines AP, BP, CP are produced to meet the circle determined by A, B and C in the points A_1, B_1 and C_1 respectively, then the triangle $A_1B_1C_1$ is equiangular with the pedal triangle of P with respect to $\triangle ABC$.

12. Prove that the angles of the pedal triangle of the orthocenter of a triangle ABC with respect to this triangle are bisected by the altitudes of $\triangle ABC$.

13. Carry through the proof of Miquel's theorem for the case that one of the points L, M and N lies on the extension of a side of the triangle, the others on the sides themselves. (Compare p. 237.)

14. Prove Miquel's theorem for the case that two of the points L, M and N are taken on the extensions of the sides and the third point on a side; also in case all three points are on the extensions.

109. The structure of a proof. We must give our attention to two other aspects of the geometrical theorems which we have been discussing, namely (1) to the structure of the proofs, (2) to the way in which the conclusions obtain validity. As to the second of these, it must inevitably occur to any thinking person to wonder why the conclusions reached in geometrical proofs are so generally regarded as incontrovertible, why we feel so sure about them, and why we expect that they will remain intact no matter what the future may bring. Let us first examine in some detail the proofs that have been given on the preceding pages.

We usually begin by drawing a diagram. Every one understands that "argument from the figure" is not valid, and that the diagram does not in itself contribute to the proof. On the other hand, the diagram is useful not only because it exhibits clearly the elements involved in a proposition, but also because in constructing it we

are forced to consider whether these elements are actually attainable.

For instance, in the statement of Theorem L (compare p. 228), there are mentioned "the perpendiculars dropped from a point P . . . upon the sides BC, CA and AB of a triangle." These phrases are meaningless unless we already know that through any point P there can always be constructed one and only one line perpendicular to a given line m. In the construction of the diagram this earlier knowledge is brought into play. From the general proposition just mentioned we conclude that the particular instance can be realized.

We look next for a combination of the elements of the given proposition which fits into and completely fills the hypothesis of a proposition we have already proved. In the proof of Theorem LIV (see p. 236), we found that we had a triangle AMN, and a point on each of its sides. This set of elements fitted the hypothesis of Theorem LIII. The conclusion of such a previously proved theorem then becomes applicable to the situation under scrutiny.

It is not difficult to see that what is involved in the drawing of a diagram is an instance of this "fitting in" process. In the example mentioned above, there are given, three times, a point and a line, viz. P and BC, P and CA, P and AB. These are the only elements in the hypothesis of the proposition which insures the possibility of dropping a perpendicular; hence we infer its conclusion in each case, i.e. the possibility of dropping perpendiculars from P to the sides of $\triangle ABC$.

The reader should have no difficulty in seeing that the entire argument used in our proofs consists of a sequence of reasonings of which the following are typical examples:

If a point P and a line m are given, there exists one and only one line passing through P and perpendicular to m. There are given a point P and the line BC, therefore there is one and only one perpendicular from P to BC.

If the line from the first to the second of three points makes a right angle with the line from the second to the third, then the line from the first to the third point is a diameter of the circle determined by the three points. The line from P (see Fig. 34, p. 230) to E makes a right angle with the line from E to D.[1] Therefore

[1] It is to be understood that P, E and D are merely short forms for designating the general concepts which they illustrate: in other words that we are not arguing from the figure.

the line PD is a diameter of the circle determined by the points P, E, D.

These arguments each consist of three parts; they are usually designated as *major premise*, *minor premise*, and *conclusion*. The argument in its entirety is called a *syllogism*. In schematic form, we have then the following formulation for the elements which make up the structure of our geometric proofs:

> *major premise:* If A holds, then B holds;
> *minor premise:* A_1 holds;
> *conclusion:* B_1 holds.

Here A_1 and B_1 may be obtained from A and B by the substitution of particular elements for some or all of the general elements entering in these.

So much for the structure of our proofs. Before passing on to other questions, it should be observed that the combination of syllogisms may lead to very complicated chains of reasoning. The analysis of such chains is one of the subjects studied in the field of the logic of mathematics. The outstanding work in that domain is the monumental *Principia Mathematica*, by A. N. Whitehead (1861–) and B. Russell (1872–). Let us now consider the validity of our conclusions.

110. The founders of the country. It should be clear that if we inquire as to the validity of a conclusion which is reached at the end of a chain of syllogisms, each of which may have several component parts, we have an experience like playing the game of "Pussy wants a corner," for the validity of the conclusion depends upon the validity of the major and minor premise [1] — and these two are of quite different character. As to the minor premise, the question is whether a certain fact A_1 actually obtains; it has to be answered by deciding whether the data have been correctly interpreted. In the examples used above (see pp. 240) it is a question of whether or not "a point P and the line BC" are given, and whether "the line from P to E makes a right angle with the line from E to D." This aspect we will leave out of consideration, assuming throughout that no errors such as misinterpretation of data have been committed. But as to the major premise, the

[1] Quite apart from this there looms up of course as a threatening spectre the question of the validity of the entire syllogistic procedure; but this is still a different question, see p. 247.

situation is quite different; it consists of a hypothesis A and a conclusion B. It can only have been attained as the conclusion in an earlier syllogism; its validity rests therefore, by the reasoning we have just gone through, upon the validity of the major premise in that earlier syllogism. And so the game is started; and once started, it appears to have a power for self-perpetuation. How long can it continue? Clearly it will go on indefinitely, unless we can find another source for validity besides that which arises from a syllogism. The situation is in some respects like that of a new resident in a country. If he is to acquire citizenship in the country there must be at least one other person there to confer this privilege upon him; and of this other person we must suppose that he is a citizen already. How did he obtain his citizenship? By having had it bestowed upon him either at his birth, or later on by some one else. But how about the first arrivals in the country? Since they were the first, there was no one there before them who could elevate them to citizenship. They must have been the founders of the country; they gave themselves citizenship, probably without asking questions as to their beliefs, their antecedents, the founders of the country, or its present rulers.

Just so, we can not escape an endless postponement of the establishment of the validity of a single conclusion in geometry, except by the creation of at least one, and probably more than one, *founder-proposition*, which acquires its right to existence from itself; or better, for which the question as to a right of existence is by common consent not raised.

These propositions are usually called the *postulates* or the *unproved propositions* of the geometry; it is upon them that the entire structure is raised by means of various combinations of syllogisms. We have now reached a partial answer to our inquiry.

(a) *Every conclusion which has validity in geometry is either a postulate or else it is obtained from the postulates by a combination of syllogisms.*

This by no means disposes of the question. Before pursuing it further we have to consider one other aspect of our subject. So far we have been concerned with the structure of geometrical reasoning. What about its contents, about the points, straight lines, circles, angles, triangles and so forth concerning which the propositions propose? What is their meaning? The obvious

answer is that these terms are introduced by definitions; e.g. two lines are called *perpendicular* if they make a right angle with each other; a circle is a line all of whose points have the same distance from some one point, etc. So far, everything is clear. But in every definition there enter a number of words (*right angle*, *distance* are some of the words entering in the definitions just quoted). How is their meaning obtained? It is obvious that we have another "pussy-wants-a-corner" game, which threatens to drag on into eternity. We cut it short by the same method that we have used before, viz. by the creation of a number of *founder-concepts*, usually called the *undefined concepts*, or the *primitive concepts*. By the side of the statement made in (a), we must therefore place a second, viz.:

(b) *Every concept admitted in geometry is either a primitive concept, or else is obtained by a definition in terms of the primitive concepts.*

A difficulty of wider scope arises here. For the definitions used in geometry contain many words besides those representing primitive concepts, or concepts "defined" in terms of them. For example, the definition of "perpendicular" on this page contains, besides the words "right angle" also the words "make," "with," "each other"; the definition of "circle" involves the troublesome words "is" and "the same." Their meaning is not covered by earlier definitions; neither have they been recognized as primitive concepts. They are borrowed from the ordinary language of everyday life. Attempts to solve this difficulty lead into the very complicated question as to the meaning of language. These are quite beyond our present purpose and must be left alone; but it is significant for our purpose to point out that a fundamental study of the foundations of geometry is thus inevitably linked up with the consideration of philosophical problems.

111. Geometry and "reality." The question of the validity of geometrical conclusions has now been resolved into two parts: (1*a*) the validity of the postulates, (2*a*) the validity of the syllogistic procedure. In a similar way the meaning of the content of geometry has acquired two aspects: (1*b*) the meaning of the primitive concepts, (2*b*) the meaning of words used in defining other concepts. For the present we leave out of consideration parts (2*a*), and (2*b*),[1] and we concentrate our attention on (1*a*) and (1*b*): the validity of the postulates and the meaning of the primitive

[1] Compare however *112*, pp. 246–8.

concepts. A great deal of study and discussion has been devoted to these questions. The outcome of it all is simply this, that with the exception of one consideration to which we shall revert immediately, *these questions are not capable of ultimate solution.* This at least is the strictly mathematical point of view; it is expressed in the significant aphorism of Bertrand Russell,[1] that "mathematics is the science in which we never know what we are talking about, nor whether what we say is true." This amounts to the statement that in any attempt to deal with these questions by mathematical methods, we would need another set of postulates and primitive concepts for which the questions would then have to be repeated. We would not be any nearer to their solution; the ultimate conclusion would merely be postponed.

Nevertheless there may be considerable gain in thus replacing one system of postulates and primitive concepts by another, namely in case the new system is simpler, contains fewer postulates or fewer undefined concepts than the earlier one; or in case the elements of the new system are more readily applicable to concrete situations. Moreover, whether such simplifications or greater adaptability be attained or not, the fact that two systems are interchangeable gives additional insight into the properties of each of them.

One further consideration is essential, viz. in order that a system be valid the postulates must not be self-contradictory, it must not be possible to deduce from them two conclusions which contradict each other. A set of postulates for geometry by means of which one could prove (a) that the sum of the angles of any triangle is equal to a straight angle, and (b) that the sum of the angles of some one triangle is less than a straight angle, could not be used. The condition which arises in this way is usually expressed by the requirement that the set of postulates must be consistent. Most readers probably have some understanding of the meaning of this requirement and it is not likely that their understanding will increase in clarity by further discussion. But it is worth while to inquire how the *consistency* of a system of postulates is demonstrated; in so doing we should also get some insight into the possibility of obtaining a basis for the validity of the postulates themselves.

The process commonly used to demonstrate the consistency of a

[1] See the *International Monthly*, Vol. 4 (1901), p. 84; Young, *Fundamental Concepts*, p. 4.

system is to adduce, or to manufacture, an "actual system" in which the postulates hold true. Any contradiction which could arise among the deductions from the postulates would then have its counterpart in this actual system. But in an "actual system" no contradictions can arise.

What however is an "actual system"? It must be some collection of entities whose existence is assured independently of the geometry with which we are concerned; it must be extraneous to the geometry. Hence the consistency of a set of postulates for geometry may become demonstrable by an appeal to a "reality" which is entirely apart from this geometry. Moreover, the holding of the postulates in such an "actual system" carries with it that a meaning has been attributed to the primitive concepts in terms of which the postulates are expressed. Thus the problem of the validity of our geometry may also find a solution by an appeal to a "reality" which is external to it. A further pursuit of this question leads once more to a fundamental philosophical problem. To deal with it satisfactorily requires extensive preliminary studies, which can not be included in our program; we must therefore leave the matter at this point.

It may help to clarify the discussion of this paragraph if we consider briefly a very simple example of a system of postulates and primitive concepts.[1]

The primitive concepts are the following: 1. A class, S, of elements a, b, c, . . .; 2. an m-class; 3. an element belongs to a class.

The postulates are the following:

I. If a and b are distinct elements which belong to S, there exists at least one m-class to which they belong.

II. If a and b are distinct elements which belong to S, there is not more than one m-class to which they both belong.

III. If two m-classes are given, there exists at least one element of S which belongs to both.

IV. There exists at least one m-class.

V. To every m-class belong at least three distinct elements of S.

VI. Not all the elements of S belong to the same m-class.

VII. There is no m-class to which more than three elements of S belong.

[1] See Veblen and Young, *Projective Geometry*, Vol. 1, p. 3; compare also J. W. Young, *Fundamental Concepts of Algebra and Geometry*, pp. 38–46.

Before taking up the consistency of this set of postulates, let us deduce some theorems from them. It is easy to see:

(a) that for any two distinct m-classes there is always exactly *one* element of S which belongs to both.

For we know by III that there is *at least* one such element; and it follows from II, that if there were more than one, the two m-classes could not be distinct.

(b) that for any two distinct elements of S there is always *exactly one* m-class to which both belong.

The reader should have no difficulty in proving this theorem.

Now to prove the consistency of the postulates, we use the following "actual system":

1. The *class* S is the set of integers 1, 2, 3, 4, 5, 6, 7. These integers are the *elements*.

2. An *m-class* is any one of the triples of integers in the following columns:

1	2	3	4	5	6	7
2	3	4	5	6	7	1
4	5	6	7	1	2	3

3. "An *element* belongs to a *class*" means that an integer occurs in a set of integers.

There should be no difficulty in verifying that all the postulates are fulfilled if the terms occurring in them are interpreted in this way.

Thus the validity of this system and of the theorems which can be established in it has been secured by the introduction of the "extraneous" integers 1, 2, . . . 7. How would *they* obtain validity?

112. The structure of mathematics. We have analyzed in some detail the structure of geometry and of its foundations, because this structure is characteristic for every domain of mathematical science, and for mathematics as a whole. The statements which have been made concerning the basis of the validity of geometrical conclusions are applicable at once to every other field of mathematics. "Any mathematical science is a body of theorems deduced from a set of axioms. A geometry is a mathematical science."[1]

Thus we can state in place of the principles of *110* (see pp. 242 and 243):

[1] See Veblen and Whitehead, *The Foundation of Differential Geometry*, p. 17.

Every conclusion which has validity in MATHEMATICS *is either a postulate or else it is obtainable from the postulates by a combination of syllogisms;* and

Every concept admitted in MATHEMATICS *is either a primitive concept or else is obtained by definition in terms of the primitive concepts.*

With respect to the validity of the postulates of mathematics and the meaning of its primitive concepts, we can only repeat the discussion of *III*, in which there was nothing that applied to geometry any more than to other parts of mathematics. But a new element enters into the situation. For, to establish consistency and validity we have to appeal now to a reality which is external to the entire domain of mathematics; actual systems like those introduced on page 246 can no longer be used. It is needless to point out that we become thereby involved in a number of philosophical problems of great difficulty. These have received considerable attention in recent years, particularly from the Viennese school. The literature on this subject consists chiefly of articles in periodicals. A simple orientation in this field can be obtained by reading a small book by H. Hahn (1879–1934), entitled *Logik, Mathematik und Naturerkennen.*

A few courageous people have undertaken to carry out a postulational treatment of mathematics in accordance with the point of view we have developed. The pioneer work in this field was done by G. Peano[1] (1858–1932); many years had to elapse before its significance was adequately recognized. It paved the way for the very extensive *Principia Mathematica* of Whitehead and Russell to which reference has been made before. Of recent date is the book by Hilbert and Bernays, *Grundlagen der Mathematik.*

Although obviously of concern to all mathematicians, this work on the foundations of the subject has, at least until very recent times, been taken up for serious consideration by only a very small number among them. I am inclined to think that this comparative neglect is due to the fact that the syllogistic procedure, which was borrowed from classical logic and had heretofore been accepted without question, had to be subjected to the same sort of analysis and to the same postulational treatment as the fields of geometry, arithmetic, etc., which were recognized as properly belonging to mathematics. The field of mathematical logic had to be created, with new symbols and with its own rather elaborate technique.

[1] *Formulario de Matematica,* Editions 1 to 5.

A deeper immersion in metaphysical questions was an unavoidable consequence; it became impossible for all but a few mathematicians to enter upon this field without sacrifice of the older interests. A special group of workers had to be prepared for these investigations. Once this was done, new developments followed each other rapidly, many new questions were brought forward. The study of the foundations of mathematics and of their philosophical significance is now well under way; it is a fruitful field and offers large scope for study and speculation to those who are well equipped both in mathematics and philosophy.[1]

Difficulties of a different kind and more puzzling in character are met by those who deal with the linguistic aspects of the problem. It is doubtful whether any real progress has been made in that direction as yet.[2] At all events, it points to a domain which lies well beyond the boundaries of our garden.

113. The foundations of geometry. We must now return to geometry and inquire to what extent the principles of *110* are adhered to in that science.

In the books on plane geometry with which the reader is familiar and which still derive in a large measure from Euclid's *Elements*,[3] a number of "axioms" and "postulates" are stated. The distinction made between these two (e.g. the following: an axiom is a statement accepted as true without proof; a postulate is a construction assumed possible[4]) is not important for our purpose; we shall continue to use the word *postulate* for both kinds of assumptions. Neither are we greatly concerned with the exact content of these postulates, since, apart from the requirement of their consistency and their simplicity (compare p. 244), the choice of postulates is arbitrary; it is usually determined by pedagogical considerations. The primitive concepts are usually not explicitly recognized. In the opening chapter of a plane geometry book we do however usually find statements like these:

"The point is the simplest geometric concept: it has position, but not magnitude."

[1] The reader may find it worth while to consult an article "Some Philosophical Aspects of Mathematics," published by the author in the *Bulletin of the American Mathematical Society*, Vol. 34, 1928, p. 438; see also the collection of essays on *The Foundations of Mathematics*, by F. P. Ramsey, published after their writer's untimely death in 1930.

[2] Attention may be called to a book by A. P. Bentley, *Linguistic Analysis of Mathematics*, and to A. Korzybski, *Science and Sanity*.

[3] Compare T. L. Heath, *The Thirteen Books of the Elements of Euclid*.

[4] See e.g., Wells and Hart, *Plane Geometry*, pp. 22, 25.

"A moving point describes a line."[1]

Among the postulates one usually finds the following:

"Two points determine a straight line."

"Two straight lines in a plane determine a point."

"A straight line-segment may be produced."

"A figure may be moved about in space with no other change than that of position, and so that any of its points may be made to coincide with any assigned point in space."

"Through a given point there can be only one parallel to a given line."

"If one circumference intersects another one, it intersects it again."[2]

These postulates are restatements and elaborations of those on which Euclid based the *Elements;* the geometry which is derived from them is therefore called *Euclidean geometry*. During the past century the foundations of this geometry have been the subject of detailed study and of far-reaching criticism. For us, the important question is whether these postulates are such that the superstructure erected on them satisfies the requirements expressed in the principles of *110* (pp. 242 and 243). The answer is *no*. That this must be the answer can easily be verified by reference to some of the proofs in the first part of this chapter. Let us take for instance the proof of Theorem L (see p. 228). To be able to conclude that the points E, D and F lie on a straight line, it is necessary, besides showing that $\angle BDF = \angle EDC$, to prove that the line-segments DE and DF lie on opposite sides of the line BC. But this has not been shown; neither is there anything in the usual postulates of Euclidean geometry which would enable us to show this. A similar question arises in the proof of Theorem LII (see p. 232). In the proofs of these theorems there have thus occurred statements which are not "either a postulate or else obtained from the postulates by a combination of syllogisms." In so far they are defective mathematical proofs; the defects do not lie in erroneous applications of the postulates, but they indicate the necessity of further postulates if the conclusion is to obtain validity. And this can indeed be done. An interesting example of a different kind but pointing to the same necessity is given by J. W. Young.[3] There

[1] See e.g., Beman and Smith, *New Plane Geometry*, pp. 2 and 3.

[2] Compare Beman and Smith, *op. cit.*, pp. 10, 24, 68; Wells and Hart, *op. cit.*, p. 50.

[3] Compare J. W. Young, *Fundamental Concepts of Algebra and Geometry*, p. 143.

it is shown that the absence of postulates concerning the order of points on a line leaves open a loophole through which the conclusion can be reached that every triangle is isosceles; it is only when such additional postulates are introduced that the conclusion fails to be valid.

The outstanding reconstruction of Euclidean geometry in accordance with the principles formulated in *110* is the one made by D. Hilbert (1862–) in his *Foundations of Geometry.*[1] There we find a clear recognition of the nature of geometry as a mathematical science. The opening sentences indicate the position: "Let us consider three distinct systems of things. The things composing the first system, we will call *points* and designate them by the letters A, B, C, . . ."; etc. "We think of these points, straight lines, and planes as having certain mutual relations, which we indicate by means of such words as 'are situated,' 'between,' 'parallel,' 'congruent,' 'continuous,' etc. The complete and exact description of these relations follows as a consequence of the *axioms of geometry.*" Thus, "point," "line," "between," "parallel," etc., are primitive concepts; in terms of them the postulational basis of Euclidean geometry is formulated.

The work of Hilbert is not by any means the only strictly mathematical basis for Euclidean geometry. Several other sets of "axioms for geometry" both preceded and followed the appearance of Hilbert's book. Among the most interesting sets is that published by Veblen in 1904,[2] in which the primitive concepts were only two in number, viz. "point" and "order." The literature on this subject has become very extensive; but there is still ample opportunity for further work. Enough has been done already to make it clear that foundations for Euclidean geometry, subject to the restrictions to which we called attention in *110* and *111*, can be established in a variety of ways.

[1] The German original, *Grundlagen der Geometrie*, was published for the first time in 1899.

[2] See O. Veblen, "A System of Axioms for Geometry," *Transactions of the American Mathematical Society*, Vol. 5, 1904, p. 343; references to other works will be found on p. 344 of this paper.

CHAPTER XI

CURVED MIRRORS

Abstruse reasoning is to the inductions of common sense what reaping is to delving. But the implements with which we reap, how are they gained? by delving. Besides, what is common sense now was abstract reasoning with earlier ages. — S. T. Coleridge, *Anima Poetae*, p. 51.

114. The proof of a postulate. Of outstanding significance among the postulates of Euclidean geometry is the "parallel postulate," stated in many textbooks in the somewhat loose form quoted on page 249: "through a given point there can be only one parallel to a given line." In Hilbert's *Foundations of Geometry*,[1] we find the following more exact formulation: "In a plane α there can be drawn through any point A, lying outside of a straight line a, one and only one straight line which does not intersect the line a. This straight line is called the parallel to a through the given point A."

In Euclid's *Elements* the postulate does not appear in this form. We do find there, however, as postulate V, the following unproved proposition:

"If a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than the two right angles."[2]

From this postulate V is then deduced a proof of Proposition 31, which asserts that "through a given point, one and only one straight line can be drawn which will be parallel to a given straight line."[3]

This postulate V of Euclid aroused the attention of the earliest commentators[4] and editors. They were impressed by the fact that

[1] *Op. cit.*, p. 11.

[2] Compare Heath, *The Thirteen Books of Euclid's Elements*, 2nd ed., Vol. I, p. 202; also R. Bonola, *Non-Euclidean Geometry*, p. 1.

[3] Compare Heath, *op. cit.*, pp. 315, 316.

[4] Most of our knowledge of the life and work of Euclid, including the *Elements*, is derived from the writings of others; compare Sanford, *op. cit.*, pp. 11 and 268–275; Heath, *op. cit.*, pp. 1–6, 202–204.

this postulate is much less direct in character than the other postulates, such as "a straight line-segment may be produced"; "two points determine a straight line," etc. It was generally thought that postulate V should be capable of proof, i.e. deducible from the earlier postulates. The history of the development of geometry for a thousand years after Euclid recounts numerous attempts at such proofs.[1] In every case, the author introduces, however, for the purpose of his proof, in addition to the earlier postulates, a new assumption which in its turn can be deduced from postulate V itself. These attempts illustrate the process

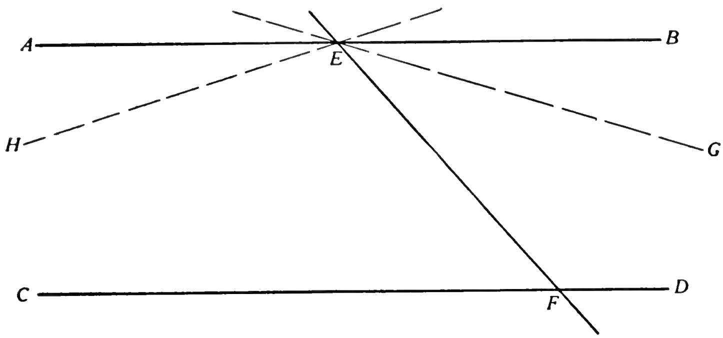


FIG. 39

mentioned in *111*. It will be instructive to consider a few of these "proofs" of postulate V.

115. A sample taken from a large collection. One of them proceeds essentially as follows:

Given. AB and CD are parallel lines, i.e. lines which do not meet, however far they may be produced; the line EF meets both AB and CD.

To prove. The sum of the angles BEF and DFE is equal to two right angles. (Compare Fig. 39.)

Proof. The sum of the two angles is either equal to two right angles, or less than that amount, or greater than that amount.

(a) Suppose $\angle BEF + \angle DFE >$ two right angles. Then we can construct through E a line EG such that $\angle GEF + \angle DFE =$ two right angles; the part of this line to the right of E will fall within the angle BEF. Now on the one hand, it follows from an earlier

[1] Compare Bonola, pp. 1–21.

proposition,[1] that EG is parallel to CD. On the other hand, the distance from the point G to the line AB will increase without limit when the distance of G from E is increased indefinitely; but since *the distance between the parallels CD and AB remains finite, the lines EG and CD must meet.* The assumption that $\angle BEF + \angle DFE >$ two right angles leads therefore to a contradiction.

(b) Suppose $\angle BEF + \angle DFE <$ two right angles. In that case $\angle AEF + \angle CFE >$ two right angles, so that the argument made in (a) can be repeated by means of a line EH, so that we shall again arrive at a contradiction.

We conclude that $\angle BEF + \angle DFE =$ two right angles.

It is now easy to prove postulate V, as follows:

Given. The lines AB and CD are met by the line EF in such a way that the sum of the angles BEF and DFE is less than two right angles.

To prove. The lines AB and CD meet on the side of B and D.

Proof. (See Fig. 39.) One of the following three possibilities must occur: (1) AB and CD are parallel; (2) they meet on the side of A and C; (3) they meet on the side of B and D.

In case (1), $\angle BEF + \angle DFE =$ two right angles as has just been proved.

In case (2), $\angle AEF + \angle CFE <$ two right angles,[2] and therefore $\angle BEF + \angle DFE >$ two right angles. Both these conclusions contradict the hypothesis and therefore the third possibility is realized.

The reasoning which we have here given is essentially that of a "proof" attributed to Proclus (410–485), the author of the most important commentary on Euclid.[3]

Let us look a little more closely at the first of these "proofs." At several points it contains statements about the relative position of lines, of whose connection with earlier postulates we may well be doubtful. We shall not discuss these statements critically because such criticisms are contrary to the spirit of Euclid's work.[4] It suffices for our present purpose to call attention to the italicized

[1] See Heath, Vol. I, p. 309, Proposition 28.

[2] See Proposition 17, Heath, Vol. I, p. 281.

[3] See Bonola, pp. 2 and 5; also Heath, Vol. I, pp. 19, 29–46; p. 207.

[4] It has been observed at an earlier point (compare pp. 248–250) that many of the arguments in ordinary Euclidean geometry need the support of additional postulates concerning the order of points and lines before they acquire validity. We do not want to complicate our discussion by pointing out these defects in every case; foregoing such criticisms keeps us nearer to the spirit of Euclid's work.

statement on page 253. It rests on the supposition that *the distance between two parallel lines remains finite.* Is this a valid supposition? In other words (compare p. 242), is it a postulate, or is it deducible from the postulates? It does not occur among the postulates I to IV, and it has not been deduced from them.[1] We can therefore only admit it as a new postulate; let us call it postulate V*a*. The proofs given above show then, if we remain within the spirit of Euclid's work, that postulate V is deducible from postulates I–IV *and Va*. And now it is interesting to show that we could equally well derive postulate V*a* from postulates I–V, so that V and V*a* are equivalent, as long as postulates I–IV hold; the argument of

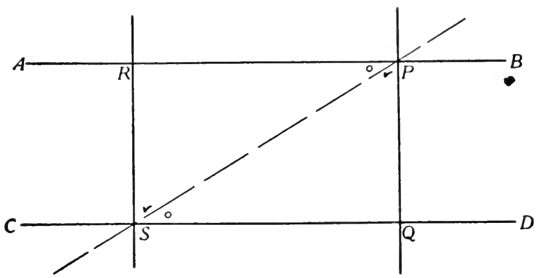


FIG. 40

Proclus is therefore not a proof of postulate V, but rather a proof of the fact that postulates V and V*a* are equivalent. We proceed to show that V*a* is a consequence of postulates I–V.

Given. AB and CD are parallel lines.

To prove. The distance between the lines AB and CD remains finite.

Proof. (See Fig. 40.) Let P be an arbitrary point on AB, and let PQ be drawn perpendicularly to AB. It follows from postulate V that $\angle BPQ + \angle DQP = 2$ right angles; and hence we conclude that $\angle DQP$ and $\angle PQC$ are right angles. If we take now a second point, R, on AB and draw RS perpendicularly to AB we find, in exactly the same way, that $\angle PRS$ and $\angle QSR$ are also right angles.

[1] Modern versions of the first four of Euclid's postulates were given on p. 249. On account of the frequent references made to them in the present chapter, a statement of them (in English translation) as occurring in Euclid's elements is given here: I. To draw a straight line from any point to any point. II. To produce a finite straight line continuously in a straight line. III. To describe a circle with any centre and distance. IV. That all right angles are equal. (Compare Heath, Vol. I, pp. 195–200.)

Moreover, it follows readily from a further application of postulate V that the alternate interior angles made by the transversal PS are also equal, so that $\angle RPS = \angle QSP$, and hence, by subtracting these equal angles from the right angles RPQ and QSR respectively, that $\angle QPS = \angle RSP$. But now the triangles RSP and PSQ have equal one side ($PS = PS$) and the two adjacent angles and are therefore congruent; consequently $RS = PQ$. Since the points P and R were chosen arbitrarily, we obtain the conclusion that two parallel lines are everywhere equally distant. The assertion of postulate V*a* is contained in this conclusion.[1]

116. An example from the 17th century. As a second example of the efforts made to get rid of postulate V, we shall look at the proposal of John Wallis (1616–1703), precursor and teacher of Newton, which consisted in replacing postulate V by another one, viz. by the postulate that "to every figure there exists a similar figure of arbitrary magnitude"[2]; we shall refer to it as postulate V*b*. The significant difference between this proposal and the work of Proclus is that the new postulate was clearly recognized as such and not tacitly assumed. The advantage of introducing it lies, according to Wallis, in its greater simplicity which secures for it more readily the assent of the reader.

By means of postulate V*b* and postulates I–IV, we can prove V as follows[3]:

Given. AB and CD are met by the transversal EF in such a way that the sum of the angles BEF and DFE is less than two right angles.

To prove. The lines AB and CD meet on the side of B and D.

Proof. (See Fig. 41.) Since $\angle PEB + \angle BEF = 2$ right angles, while $\angle DFE + \angle BEF < 2$ right angles, it follows that $\angle PEB > \angle EFD$. Therefore, if the line FE, with FD attached to it, be moved along the line QP until F coincides with E, the line FD will fall inside the angle BEP, i.e. in the position EG; and in the course of the motion, the line FD will occupy a position E_1G_1 in which it cuts the line EB in G_1. Consider now the triangle

[1] In the arguments made on this and preceding pages, we have drawn on the reader's knowledge of the early parts of Plane Geometry, in particular upon his knowledge of Postulates I to IV and of the theorems which can be deduced from them. It will be worth while to verify the validity of these arguments by reference to Heath's edition of the *Elements*.

[2] See Bonola, pp. 15–17; Heath, Vol. I, p. 210.

[3] Compare the references given in footnote 2.

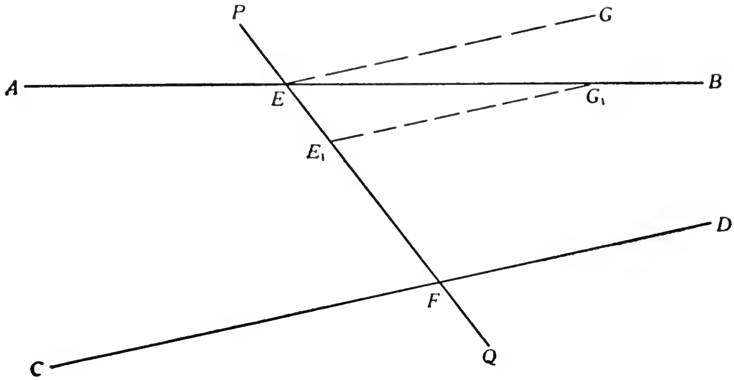


FIG. 41

EE_1G_1 and apply postulate V*b* to it; it will then be possible to construct another triangle similar to triangle EE_1G_1, in which EF corresponds to the side EE_1. It follows then that the other sides of this new triangle must fall along EB and FD, so that the lines AB and CD actually meet on the side of B and D.

The reader will have little difficulty, while admiring Wallis' ingenuity, in pointing out the weak points in his proof from the point of view of modern criticism; but we have agreed not to pursue such criticisms (compare p. 253, footnote 4). However, as shown in the next proposition, we can deduce Wallis' postulate V*b* from postulates I–V; therefore the situation is entirely analogous to the one for postulate V*a*, viz. under acceptance of I to IV, postulates V and V*b* are equivalent.

The use of the word "figure" introduces into postulate V*b* an element of vagueness; in the proof which was reproduced above, the postulate was only used for a triangle. We shall therefore prove V*b* for this case:

Given. A triangle ABC and a line segment PQ.

To prove. There exists a triangle similar to ABC in which PQ corresponds to AB.

Proof. (See Fig. 42.) We can construct lines PS and QT such that $\angle SPQ = \angle CAB$, and $\angle TQP = \angle CBA$.[1] Since the sum of the angles CAB and CBA is less than 2 right angles (Heath, p. 281, Proposition 17), the same can be said about the angles SPQ and

[1] Compare Heath, Vol. I, p. 294, Proposition 23.

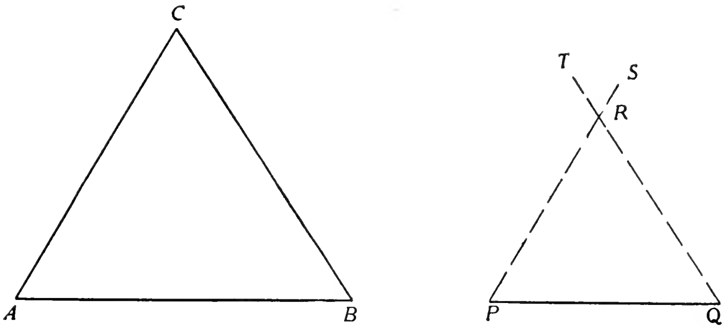


FIG. 42

TQP; consequently, we conclude by use of postulate V that *SP* and *TQ* meet — let us call their common point *R*. The triangles *ABC* and *PQR* have then two pairs of equal angles, and therefore three pairs of equal angles [1]; hence they are similar triangles.

117. The weight of authority. While the numerous attempts to "prove" the parallel postulate failed in their main objective, they were very useful in bringing out important connections between a number of theorems. It is of particular interest that this work led to a recognition of the fact that the theorem which asserts that the sum of the angles of a triangle is equal to two right angles is equivalent (in the same sense in which this term was used in the preceding sections) to the parallel postulate.

It is probably difficult for us to realize the enormous hold which the "truths" of geometry had upon the scientists of the 17th and 18th centuries. That any of the conclusions of the geometry of Euclid should fail to be "true" was to most of them inconceivable. This was due, in part, to the fact that the abstract character of geometry was but imperfectly understood; and partly to the weight of authority which made any departure from recognized and established truths a dangerous undertaking.[2] If the idea that the postulates of Euclidean geometry might not express "absolute truths" had come to *Girolamo Saccheri* (1667–1733), he would

[1] Compare Heath, Vol. I, p. 316, Proposition 32. The connection between postulate V and the angle sum of the triangle is dealt with in subsequent sections.

[2] Especially important in this connection are the views of the philosopher Immanuel Kant (1724–1804), who held that the conception of space as developed in Euclidean geometry was an "a priori truth" inherent in the human mind.

probably have forestalled by over a century the discovery of non-Euclidean geometry. For he considered a quadrilateral $ABCD$ (see Fig. 43) of which the sides AB and CD are equal, and the angles at B and at C are right angles. After having shown on the basis of postulates I to IV that the angles at A and at D must be *equal* (see *118*, 1), he admits the possibility of *three* hypotheses, viz. (1) that these angles are both right angles, (2) that they are both obtuse, (3) that they are both acute. The first of these is equivalent to the parallel postulate; Saccheri hoped to *prove* this postulate by showing that the other two possibilities lead to contradictions. His method is perfect; but his hope was not realized, for no contradictions revealed themselves.[1] Had he gone a step further and drawn the conclusion, very bold for his time, that each of his hypotheses can form the basis for a geometry, he would have been the founder, instead of a forerunner, of non-Euclidean geometry.

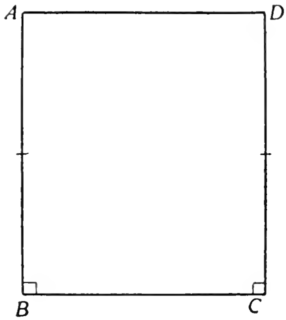


FIG. 43

It will be of interest to follow Saccheri's argument for a little distance. His work is contained in a book published in Milan, in 1733, with the title "Euclides ab omni naevo vindicatus: sive conatus geometricus quo stabiluntur prima ipsa universae geometriae principia." The three possibilities mentioned above are referred to by him as the hypothesis of the right angle, the hypothesis of the obtuse angle, and the hypothesis of the acute angle; we shall call them Hyp. I, II and III respectively. Moreover, we shall, for convenience of description, use the phrase "*rectangle of type a*" for a quadrilateral like the one in Fig. 43, in which two opposite sides are equal and in which two adjacent angles, having these equal segments on their non-common side, are right angles.

Theorem LV. If for any rectangle of type a in the plane hypothesis I holds, then it holds for every such rectangle in the plane.

Proof.[2] Let us suppose that rectangle $ABCD$ (see Fig. 44) is of type a, and that Hyp. I holds for it.

[1] Compare Bonola, pp. 22–44; Heath, p. 211.

[2] In this and in subsequent proofs we shall draw upon some facts which are established in *118*.

(a) We consider first a rectangle A_1BCD_1, of type a, obtained by laying off, on the equal sides BA and CD, equal segments BA_1 and CD_1 shorter than the sides BA and CD themselves. Since Hyp. I holds for $ABCD$, we know that $AD = BC$ (see *118*, 4); moreover angles A and D being right angles, and the segments AA_1 and DD_1 being equal, it follows that AA_1D_1D is also of type a.

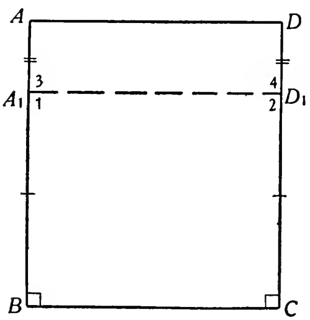


FIG. 44

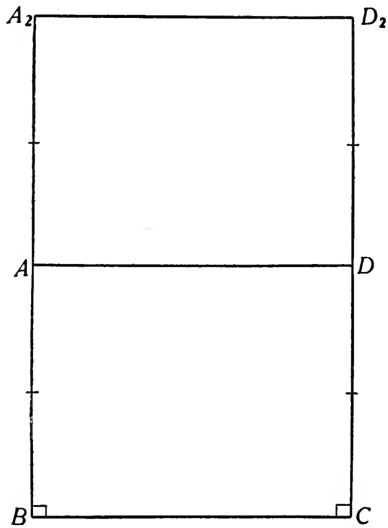


FIG. 45

Suppose now that for BCD_1A_1, Hyp. I did not hold; then either II or III must hold for it. In the former case, the angles 1 and 2 would be obtuse, and therefore angles 3 and 4 acute; from the first of these facts, we would conclude that $A_1D_1 < BC$, while from the second would follow that $A_1D_1 > AD$ (by use of *118*, 4). But since $AD = BC$, these two conclusions contradict each other; therefore Hypothesis II can not hold for BCD_1A_1. A similar argument (see *118*, 5) shows that Hypothesis III can not hold for it. Consequently, if I holds for $ABCD$, it must also hold for A_1BCD_1.

(b) Take next a rectangle of type a obtained by extending the segments BA and CD to points A_2 and D_2 respectively (see Fig. 45). If $BA_2 = 2 \cdot BA$ and $CD_2 = 2 \cdot CD$, and if Hyp. I holds for $ABCD$ it is easy to prove by means of congruent triangles (see *118*, 6) that the same hypothesis must also hold for A_2BCD_2. If $BA_2 = k \cdot BA$ and $CD_2 = k \cdot CD$ for an arbitrary natural number k, a repetition of the argument and the use of mathematical induction

will lead to the same conclusion. If BA_2 and CD_2 are equal, but not integral, multiples of BA and CD respectively (Fig. 46), we can determine an integer k, such that $k \cdot BA < BA_2 < (k + 1) \cdot BA$, and $k \cdot CD < CD_2 < (k + 1) \cdot CD$[1]; suppose that $BA_3 = k \cdot BA$, $BA_4 = (k + 1) \cdot BA$, and that $CD_3 = k \cdot CD$, $CD_4 = (k + 1) \cdot CD$. The discussion just preceding shows then that Hyp. I holds for BA_4D_4C and from this fact we conclude, by applying to the latter rectangle the argument of (a), that it holds for BA_2D_2C. Thus we have shown that if I holds for $ABCD$, it also holds for A_2BCD_2, no matter how A_2 and D_2 are selected on the extensions of BA and CD respectively, provided $BA_2 = CD_2$, so that A_2BCD_2 is a rectangle of type a.

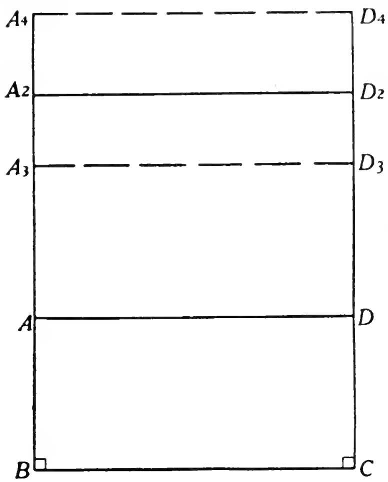


Fig. 46

(c) It remains to show that if Hyp. I holds for $ABCD$ it will also hold for a rectangle of type a whose base (i.e. the side of the rectangle which lies on the common side of the adjacent right angles) is not equal to BC, as e.g. the quadrilateral $A'B'C'D'$ whose base is $B'C'$ (see Fig. 47). To do this, we construct an auxiliary rectangle $A_1B_1C_1D_1$ of type a, in which the base B_1C_1 is equal to BC and whose side B_1A_1 is equal to $B'C'$. From what we have proved so far, it follows that if Hyp. I holds in either of two rectangles of type a with equal base, then it holds in the other as well. In particular, we know that Hyp. I holds in this auxiliary rectangle; hence A_1 is a right angle. Comparing now $A_1B_1C_1D_1$ with $A'B'C'D'$, looking upon A_1B_1 and $B'C'$ as their bases, the

[1] That it is possible to determine such an integer k appeals to our intuitive concepts concerning magnitudes and is readily verified if the ratio of BA_2 to BA is a rational number. That such a number always exists is an assumption which has been recognized as such only in modern times in the more critical study of the foundations of geometry. It is explicitly stated among the postulates of geometry of Hilbert (see Hilbert, *Foundations of Geometry*, p. 25) and is usually referred to as the *axiom of Archimedes*. Compare Bonola, pp. 29 and 30; see also p. 188.

preceding argument becomes again applicable, so that Hyp. I must hold for $A'B'C'D'$. Thus the proof of our theorem is complete.

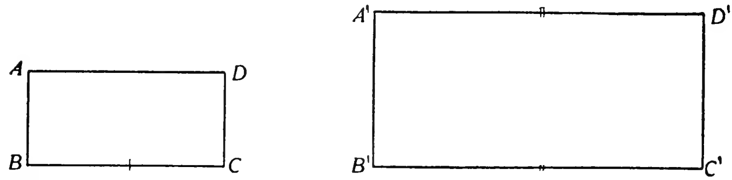
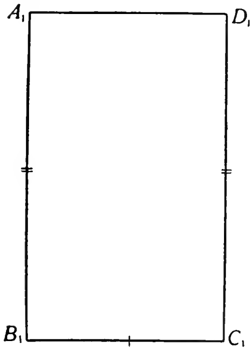


FIG. 47



Saccheri proves next corresponding theorems for Hypotheses II and III [1]; in these proofs arguments are used of which some can be admitted only provided we agree to remain within the spirit of Euclid's work. These theorems together with Theorem LV make it clear that the Hypotheses I, II and III are not hypotheses concerning a particular rectangle but rather properties of the entire plane in which the rectangle lies. Thus we are led to conceive the existence of three kinds of planes, viz. (a) such planes that for every rectangle of type *a* within it Hypothesis I holds, (b) such in which every rectangle of type *a* satisfies Hypothesis II, and (c) such in which every rectangle of type *a* satisfies Hypothesis III. A plane can then be characterized by the statement as to which of the three hypotheses holds for its rectangles of type *a*.

We must follow Saccheri's argument a little further. It is shown next that, according as Hypothesis I, II or III holds for a plane, the sum of the angles of a triangle is equal to, greater than or less than two right angles.

In the first place we observe that these conclusions will hold for any triangle, provided they hold for triangles which have one right angle. For if in triangle ABC (see Fig. 48), we draw $AD \perp BC$, we obtain the two right-angled triangles ABD and ACD. If we denote the angle sums of triangles ABC, I and II, expressed in radians (see p. 121), by S, S_1 and S_2 respectively, we see at once

[1] See e.g. Bonola, pp. 26–28.

that $S = S_1 + S_2 - \pi$.[1] Hence if $S_1 = S_2 = \pi$, then $S = \pi$; if $S_1 > \pi$ and $S_2 > \pi$, then $S > \pi$; and if $S_1 < \pi$ and $S_2 < \pi$, then $S < \pi$.

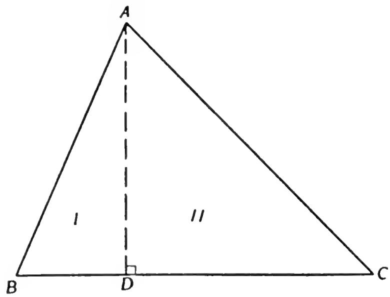


FIG. 48

Theorem LVI. According as Hypotheses I, II or III hold in a plane the angle sum S of a triangle in this plane is equal to, greater than or less than π.

Proof. In accordance with the discussion preceding the statement of the theorem, we now limit ourselves to the consideration of right-angled triangles. Suppose then that ABC is a triangle with a right angle at B. If we draw CD perpendicular to CB and equal to BA, we obtain in $ABCD$ a rectangle of type a. If Hyp. I holds, it is easily shown that the triangles ABC and ADC are congruent (see *118*, 7) so that $\angle BAC = \angle ACD$. In this case we find that the angle sum S of triangle ABC is equal to $\angle B + \angle BCA + \angle ACD = \pi$. If Hypothesis II holds, the angles CDA and DAB are obtuse, and $DA < CB$. Consequently, in triangles ABC and ADC, two sides are equal ($AC = AC$, $AB = CD$) while the third side is longer in the first than in the second. From this it follows by means of a well-known proposition[2], that $\angle BAC > \angle ACD$. Therefore

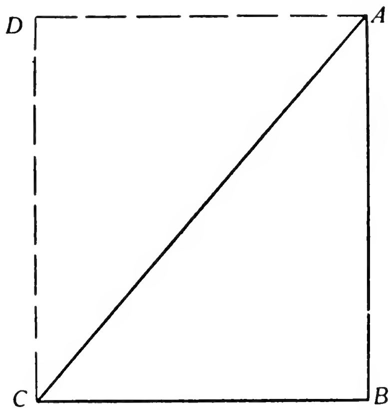


FIG. 49

$$S = \angle B + \angle BCA + \angle BAC > \angle B + \angle BCA + \angle ACD$$
$$= \angle B + \angle BCD = \pi.$$

[1] If $\triangle ABC$ has an obtuse angle, we obtain the same result, provided the perpendicular is drawn through the vertex of the obtuse angle.

[2] See Heath, Vol. I, p. 299, Proposition 25.

In case Hyp. III holds, the same argument shows that $S < \pi$ (see *118*, 8).

Corollary. According as the angle sum S of any one triangle is equal to, greater than or less than two right angles, Hypotheses I, II or III must hold.

The proof of this corollary is left to the reader (see *118*, 9). In conjunction with Theorem LVI, it shows that the hypotheses concerning the angle sum of a triangle are equivalent to Hypotheses I, II, III in the sense discussed on page 254.

We shall not pursue the work of Saccheri beyond this point. His method consists in formulating various propositions equivalent to Hypotheses I, II, III so as to obtain the conclusion that II and III lead to "absurdities." Now, the entire force of the argument depends upon what we mean by an "absurdity." In the non-mathematical sense it means anything which violates a firm conviction, or a prejudice, or a deeply-rooted intuitive certainty, or an observation in nature. Because convictions, prejudices and "certainties" vary a good deal from individual to individual, and because observations in nature are frequently dependent upon the observer and moreover subject to inaccuracies, it happens very often that what is considered as "absurd" by one person is acceptable to another. In the mathematical sense a conclusion is absurd only if it is in direct contradiction with another conclusion which has been recognized as *valid* (compare pp. 241, 242 and 247). For our problem this would mean direct contradiction with the earlier postulates of geometry or with theorems proved by means of them. On this test none of Saccheri's conclusions are absurd however much they seem to go counter to primitive intuitive ideas concerning space. Hence his effort to clear Euclid of all "blemish" did not succeed.[1]

The history of the subject presents many other noteworthy attempts to prove the parallel postulate; they are instructive and interesting examples of the resistance of the human mind to radical innovations. But we must now pass on to a consideration of the work of some of those who broke away from the old moorings. Some further study of the matters that have been discussed will be useful.

[1] The interested reader will find ample opportunity in the literature already quoted for acquainting himself in greater detail with Saccheri's deductions and with other work of a similar nature, to which only brief reference has been made in the text.

118. Working under new regulations.[1]

1. Prove that if in a quadrilateral $ABCD$ the opposite sides AB and DC are equal and the adjacent angles B and C are right angles, then the angles A and D are equal (see Fig. 43, p. 258).

2. Prove that an exterior angle of a triangle is always greater than either of the non-adjacent interior angles.

3. Prove that if in a quadrilateral $ABCD$ (see Fig. 43) the adjacent angles B and C are right angles while the opposite sides AB and DC are unequal, then the angles A and D are unequal and the larger one of these angles is adjacent to the shorter one of the sides AB and DC.

4. Prove that if, with the hypotheses of 1, the angles A and D are right angles, then $AD = BC$; if they are obtuse then $AD < BC$, and if they are acute then $AD > BC$.

Hint. Draw a line perpendicular to BC at its midpoint.

5. Show that if Hypothesis I holds for $ABCD$ (see Fig. 44, p. 259), then Hypothesis III can not hold for A_1BCD_1.

6. (See Fig. 45, p. 259.) Prove that if $BA_2 = 2 \cdot BA$ and $CD_2 = 2 \cdot CD$, then Hyp. I holds for A_2BCD_2, provided it holds for $ABCD$ (compare p. 259 (b)).

7. Prove that if Hypothesis I holds (see Fig. 49, p. 262), then triangles ABC and ADC are congruent.

8. Prove that if Hypothesis III holds, then the angle sum of a right triangle is less than two right angles (compare p. 263).

9. Prove the corollary of Theorem LVI stated on p. 263.

119. The beginning of a new geometry. Among those who had occupied themselves with the parallel postulate was the Hungarian mathematician Wolfgang Bolyai (1775–1856),[2] a friend of the great German mathematician Carl Friedrich Gauss (1777–1855), with whom he kept up correspondence after their student days in Göttingen. Gauss continued to live there as professor of astronomy from 1807 until his death; during that time he made contributions of fundamental character to every field of mathematics and its applications. Both Bolyai and Gauss had been interested since their youth in the problem of parallels. The former had published a number of papers on the subject. But his chief claim to fame rests on a work published in 1832–33 with the title "Tentamen juventutem studiosam in elementa Matheseos purae, elementaris ac sublimioris, methodo intuitiva, evidentiaque huic propria,

[1] Only postulates I–IV and their deductions are to be used in proving the following propositions.

[2] His full name is Wolfgang Bolyai de Bolya.

introducendi"; and it is particularly the appendix to the first volume of this work for which it is remembered. This appendix was written by the author's son, Johann Bolyai, and has the title "Appendix Scientiam spatii absolute veram exhibens"; it is one of the sources from which Non-Euclidean geometry developed. Johann Bolyai (1802–1860), barely thirty years old when the "Tentamen" with its famous appendix was published, was at that time an officer in the Hungarian army.[1] Introduced by his father to the problem of parallels, he had been occupied with this question for many years before the publication of his work, as is evident from letters he wrote to his father as early as 1823. In his "absolute theory" of space he built up a geometry entirely independent of Euclid's postulate of parallels, thus demonstrating the possibility of a geometry in which this postulate does not hold and consequently solving the old problem by showing that the battle-worn fifth postulate was not provable by means of the other postulates of Euclid.

When Wolfgang Bolyai sent a copy of his son's work to his old friend Gauss, he received the following reply: "If I commenced by saying that I am unable to praise this work, you would certainly be surprised for a moment. But I cannot say otherwise. To praise it, would be to praise myself. Indeed the whole contents of the work, the path taken by your son, the results to which he is led, coincide almost entirely with my meditations, which have occupied my mind partly for the last thirty or thirty-five years. . . . So far as my own work is concerned, of which up till now I have put little on paper, my intention was not to let it be published during my lifetime. . . . I have found very few people who could regard with any special interest what I communicated to them on this subject." From other letters of Gauss, it becomes clear that as early as 1816 he had projected what he then called Anti-Euclidean geometry. Later he called it Astral geometry, and finally Non-Euclidean geometry. But he did not allow his ideas to become known outside a small circle of trusted friends for fear of being misunderstood. We have in this attitude of Gauss, even more strongly than in the work of Saccheri, an example of the powerful hold accepted ideas on fundamental questions have on the minds of men, and of the fact that it takes unusual courage combined with profound insight to strike out into new paths.

[1] Compare Cajori, *History of Mathematics*, p. 302.

Before proceeding with an indication of the new theory, we must mention another aspect of its remarkable history. Beginning in 1829, a number of papers were published by the Russian mathematician Nicolai Ivanovitch Lobatschewsky (1793–1856), professor of mathematics at the University of Kasan, in which the author made known his work on a new geometry, more general than that of Euclid, in which two parallels to a given line can be drawn through a point outside that line. As it turned out this new geometry was equivalent to the systems developed by Gauss and by Bolyai. A great deal of study has been devoted to the relations between the work of these three men and to the possibilities of influences of one of them upon the others; this study has led to the conviction that all of them should be recognized as founders of Non-Euclidean geometry. For an account of these historical investigations the reader is referred to the book of Bonola, of which we have already made frequent use and in which most of the original treatises on the subject are mentioned; of great value is the treatise by Engel and Stäckel, *Theorie der Parallellinien von Euclid bis auf Gauss*, published in 1895.

120. The geometry of Lobatschewsky. We must now try to get some idea of the new geometry whose birth was such a long and painful process. We shall develop for this purpose a few of the early theorems of Non-Euclidean geometry, as worked out in Lobatschewsky's *Geometrical Researches on the Theory of Parallels*.[1]

A number of theorems which depend only upon postulates I to IV are used by Lobatschewsky in carrying out his theory. Of these the following will suffice for our purpose; they will be stated here because they have to be referred to repeatedly in the present section and the next.

I. Two straight lines can not intersect in two points.

II. A straight line sufficiently produced both ways must go out beyond all bounds, and in such way cuts a bounded plane into two parts.

III. Vertical angles, where the sides of one are productions of the sides of the other, are equal.

IV. Two straight lines can not intersect, if a third cuts them at the same angle.

V. In a rectilineal triangle, a greater side lies opposite a greater

[1] A translation of this work by G. B. Halsted was published in 1891; it was republished in 1914 by the Open Court Publishing Co.

angle. In a right-angled triangle, the hypothenuse is greater than either of the other sides, and the two angles adjacent to it are acute.

VI. Rectilineal triangles are congruent if they have a side and two angles equal, or two sides and the included angle equal, or two sides and the angle opposite the greater equal, or three sides equal.

We introduce now the concept of *parallel lines*. We draw through the point P outside the line l the perpendicular PA (compare *122*, 1), and also the line m, perpendicular to PA (see Fig. 50); we shall designate by a the length of AP, measured in some unit which we may choose at pleasure. Then it follows from IV, that the lines l and m will not intersect. Draw now various "half" lines starting at P and entering the angle MPA; some of them, like PA,

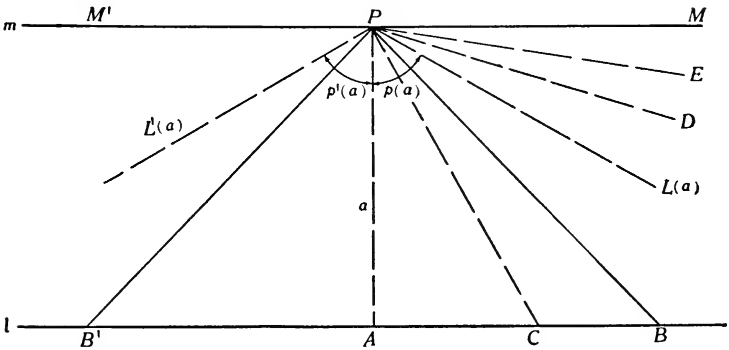


FIG. 50

intersect l, others, like m, do not intersect l. Those of the first kind we call *cutting lines;* those of the second kind, *non-cutting lines.* If PB is a cutting line, then any other line, for which, as for PC, $\angle APC < \angle APB$ will also be a cutting line. For PC must divide the bounded part of the plane which is formed by $\triangle PAB$ into two parts, in virtue of II. But it can not intersect either PA or PB in any other point besides P, in accordance with I; therefore PC must intersect AB, i.e. it must be a cutting line. On the other hand, if PD is a non-cutting line, then any other line, for which, as for PE, $\angle APE > \angle APD$, must also be a non-cutting line. For, if PE were a cutting line, it would follow by the argument we have just made that PD is also a cutting line. If we think now of a line rotating continuously about P, starting from the position PA, to the position PM, we pass from cutting lines to non-cutting lines.

Let us now think of the angles which these various lines make with PA; they are measured, in terms of any unit we wish to choose (for the sake of definiteness we will take the radian as a unit, so that the angle APM is measured by the real number $\frac{\pi}{2}$), by positive real numbers. Let α denote the measure of the angle which PA makes with any cutting line, and β the measure of the angle made with any non-cutting line. Let us now divide the system of rational numbers into two sets A and B, by putting in A all negative rational numbers and all positive rational numbers less than or equal to any α; in B all rational numbers equal to or greater than any β and all positive rational numbers which exceed $\frac{\pi}{2}$. This division constitutes a *cut* (A, B) (compare Definition XI, page 60; compare also *39*, 8). Consequently there is a real number which measures an angle, such that any positive angle less than it corresponds to a cutting line, while any angle greater than it corresponds to a non-cutting line. This angle may very well change as the distance a changes. We shall therefore designate it by $p(a)$ and we shall call it the angle of parallelism for a on the right. The half-line corresponding to $p(a)$ we shall denote by $L(a)$ and it will be called the right parallel through P to l. By an exactly analogous discussion, we reach the concept of the angle of parallelism for a on the left; we designate it by $p'(a)$. The corresponding half-line through P will be called the left parallel through P to l and will be denoted by $L'(a)$. The contents of this paragraph are summarized as follows:

Definition XXXIX. The parallels to the line l through a point P at a distance a from l are the lines which separate the cutting lines through P from the non-cutting lines; they are denoted by $L(a)$ and $L'(a)$.

Definition XL. The angles of parallelism for a point P at a distance a from the line l are the angles formed by the perpendicular from P to l with the parallels to l through P; they are denoted by $p(a)$ and $p'(a)$.

This is the concept of parallelism in Lobatschewsky's geometry. Since $M'PM$ is surely a non-cutting line, it follows that neither $p(a)$ nor $p'(a)$ can exceed $\frac{\pi}{2}$; the only possibilities are therefore

contained in the inequalities $p(a) \leqq \frac{\pi}{2}$ and $p'(a) \leqq \frac{\pi}{2}\cdot$ It is shown without difficulty that $p(a) = p'(a)$ (see *122*, 3). Hence if $p(a) = \frac{\pi}{2}$, then $p'(a) = \frac{\pi}{2}$ also, so that the half-lines $L(a)$ and $L'(a)$ are parts of one straight line. In this case we have a single parallel to l through P; if this happens for every point P and every line l which does not contain P, we have a Euclidean geometry. But if $p(a) = p'(a) < \frac{\pi}{2}$, we have a different state of affairs. There are then two parallels to l through P, one on each side of PA; they are not parts of one straight line, but meet at P under an angle

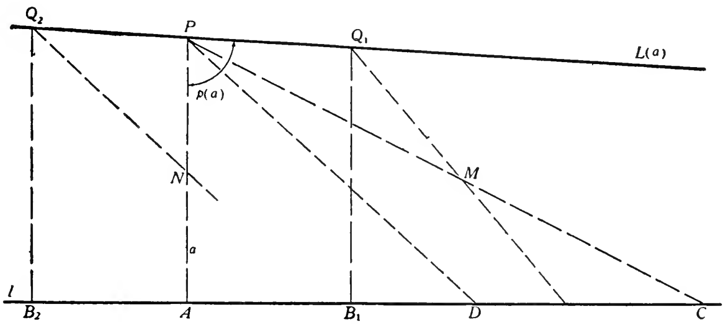


FIG. 51

smaller than π. If *this* happens for every point P and every line l which does not contain P, we have a geometry of Lobatschewsky.[1]

121. Plane geometry in a new form. On the basis of those definitions, we can now develop a number of interesting theorems.[2]

Theorem LVII. If Q is an arbitrary point on the right parallel $L(a)$ through P to l, then $L(a)$ is also the right parallel through Q to l.

Proof. (See Fig. 51.) (a) Let Q_1 be a point on $L(a)$ to the right of P. From Q_1 we draw the line Q_1B_1 perpendicular to l. Clearly $L(a)$ is not a cutting line from Q_1 to l. To prove that it is the right parallel from Q_1 to l, it is then only necessary to show that every

[1] See the paragraph preceding Theorem LXIII.

[2] Although Lobatschewsky's treatment of the following propositions has been simplified in several respects by later writers, we adhere quite closely to the original version.

line through Q_1 which enters into the angle formed by B_1Q_1 and $L(a)$ is a cutting line for Q_1. For this purpose, we connect an arbitrary point M within this angle with P, so as to obtain the line PM. Since $L(a)$ is the parallel to l through P, the line PM is a cutting line; we designate by C the point in which it meets l. Thus we obtain the triangle PAC, of which the side PC is met by Q_1M at M. It follows from I and II that the line Q_1M must cut either the side PA or the side AC of $\triangle PAC$. It can not cut PA because Q_1M and Q_1B_1 are on the same side of PA [1]; hence it must meet AC and is therefore a cutting line. This completes the proof for the point Q_1.

(b) Take now a point Q_2 on $L(a)$ on the other side of P and drop the perpendicular Q_2B_2 from Q to l. To prove that $L(a)$ is the right parallel to l through Q_2 we have again to show that every line through Q_2 which lies within the angle B_2Q_2P is a cutting line; suppose that Q_2N is such a line. Since it enters the area bounded by the quadrilateral APQ_2B_2, it must cut the boundary line of this quadrilateral in at least two points, in accordance with II. Since it meets PQ_2 and Q_2B_2 in Q_2, it can not meet them again, by I, so that it must meet either B_2A or AP. In the former case the line Q_2N is evidently a cutting line, so that the proposition is proved. In the latter case let N be the point of intersection of the line with AP; draw then the line PD on the same side of $L(a)$ as Q_2N, and so that angles NQ_2P and DPQ_1 are equal. This line PD is then a cutting line; let its intersection with l be the point D. The line Q_2N enters the triangle by APD at N and must therefore meet its bounding lines a second time. This can not take place on AP; neither can it take place on PD, since it follows from IV, that the lines Q_2N and PD do not intersect. Consequently Q_2N must, if drawn out far enough, meet AD and therefore be a cutting line.

Corollary. If Q is an arbitrary point on the left parallel $L'(a)$ to l through P, then $L'(a)$ is also the left parallel to l through Q.

The proof of this corollary proceeds like the proof of Theorem LVII.

In describing the lines $L(a)$ and $L'(a)$ we can now omit mention of the point P. There is now definite meaning in a statement like "l_1 is right parallel to l"; thus we are prepared to consider the following theorem:

[1] The reader will recognize the weakness of the argument at this point; the critique which it requires falls outside the spirit of Euclid's work and outside the scope of our present interest. (Compare p. 253.)

Theorem LVIII. If l_1 is right parallel to l, then l is right parallel to l_1.

Proof. (See Fig. 52.) From an arbitrary point P_1 on l_1 draw the line P_1P perpendicular to l at P; from P, draw the perpendicular PP_2 to l_1. The lines l and l_1 certainly do not intersect. To prove that l is right parallel to l_1, it will therefore suffice to show that any line through P which, like PB, lies within the angle APP_2 will intersect l_1. We draw the line from P_1 perpendicular to PB at Q. Since P_1QP is a right angle and P_1PQ less than a right angle, it follows from V that $P_1P > P_1Q$. We can therefore determine on P_1P a point Q_1 such that $P_1Q_1 = P_1Q$. Through Q_1 we draw a

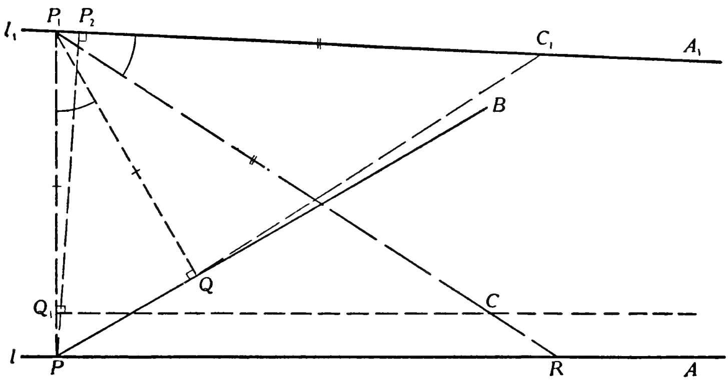


FIG. 52

line Q_1C perpendicular to P_1P; it follows then from IV that Q_1C and PA do not intersect. Furthermore we draw through P_1 the line P_1R such that $\angle A_1P_1R = \angle QP_1Q_1$. The line P_1R is then clearly a cutting line for l, since l_1 is parallel to l (compare Definition XXXIX); let R be its point of intersection with l. We have then the triangle P_1PR, into whose area Q_1C enters at Q_1; the line Q_1C must therefore meet another side of this triangle. Since it does not intersect l it must meet P_1R; let C be the point of intersection. We obtain in this way a triangle P_1Q_1C. Now we lay off on P_1A_1 a segment P_1C_1 equal to P_1C; by connecting C_1 and Q, we obtain the triangle P_1QC_1. In the triangles P_1Q_1C and P_1QC_1, we have now $P_1Q_1 = P_1Q$, $P_1C = P_1C_1$ and $\angle Q_1P_1C = \angle QP_1C_1$ (for $\angle Q_1P_1C = \angle Q_1P_1Q + \angle QP_1C = \angle A_1P_1C + \angle QP_1C = \angle A_1P_1Q$).

It follows therefore from VI that these two triangles are congruent, so that $\angle P_1QC_1 = \angle P_1Q_1C$, which is a right angle. We conclude that QC_1 must lie along the line PQB (compare *122*, 1); in other words that PB intersects l_1.

Corollary. If l_1 is left parallel to l, then l is left parallel to l_1.

The proof of this corollary is reserved for the reader (see *122*, 7).

Theorem LIX. The angle sum of a triangle can not exceed two right angles.

Proof. Suppose there were a triangle ABC, whose angle sum S were greater than π; we would then have $S = \pi + a$, where $a > 0$.

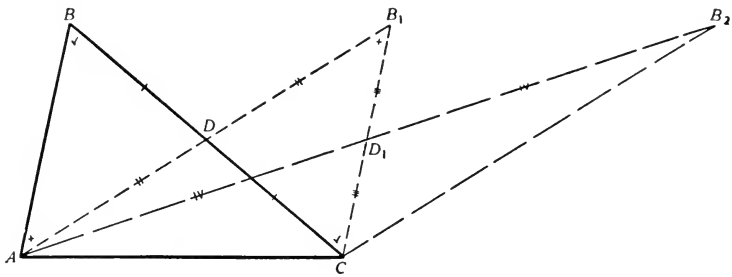


Fig. 53

Draw then the line through A and through the middle point D of the side BC (see Fig. 53), and determine on this line a point B_1 such that $AD = DB_1$; connect C and B_1 by a straight line. In virtue of III, the angles BDA and B_1DC are equal. Therefore, by VI, the triangles ABD and B_1DC are congruent, so that $\angle DCB_1 = \angle ABD$ and $\angle DB_1C = \angle DAB$. In the new triangle AB_1C, the angles B_1AC and B_1 have then a sum equal to $\angle BAC$, whose magnitude we shall denote by A, while the angle ACB_1 is equal to the sum of the other two angles of $\triangle ABC$. Therefore the angle sum S_1 of $\triangle AB_1C$ is equal to S and hence to $\pi + a$. For,

$$\begin{aligned} S_1 &= \angle CAB_1 + \angle AB_1C + \angle B_1CA \\ &= \angle CAB + \angle B_1CA \\ &= \angle CAB + \angle B_1CB + \angle BCA \\ &= \angle CAB + \angle CBA + \angle BCA = S. \end{aligned}$$

Moreover, since the sum of the angles B_1AC and B_1 is equal to A, one of them must be less than, or at most equal to, $\frac{A}{2}\cdot$ Let us sup-

pose $\angle B_1AC$ to be that one, so that $\angle B_1AC \leqq \frac{A}{2}$. We join the vertex A of this angle to the midpoint D_1 of B_1C, extend the line to B_2 so as to make D_1B_2 equal to AD_1, and complete the triangle AB_2C. It is easy to show that the angle sum S_2 of $\triangle AB_2C$ is then equal to that of $\triangle AB_1C$, so that $S_2 = S_1 = S = \pi + a$; also that the sum of the two angles B_2AC and B_2 is equal to $\angle B_1AC$, and hence less than or equal to $\frac{A}{2}$. Moreover, one of these two angles is then at most equal to $\frac{1}{2}$ of $\angle B_1AC$; supposing the angle B_2 to be that one, we have $B_2 \leqq \frac{1}{2}\angle B_1AC \leqq \frac{A}{4}$. Now we repeat the process, starting with a line from B_2 to the midpoint of the side opposite it in $\triangle B_2AC$. Continuing in this manner, we would obtain a sequence of triangles $T_1, T_2, \ldots T_k, \ldots$, in each of which the angle sum is equal to $\pi + a$; moreover, there would be a sequence of angles, one in each triangle of the sequence $T_1, T_2, \ldots T_k, \ldots$, whose magnitudes are successively less than or equal to $\frac{A}{2}, \frac{A}{4}, \frac{A}{8}, \ldots \frac{A}{2^k}, \ldots$. If we take k large enough to make $2^ka > A$, the triangle T_k would have an angle sum equal to $\pi + a$, and it would contain an angle whose magnitude is less than a. To this triangle we would apply the same process once more, and then we would finally reach serious trouble. For this last step would reveal a triangle whose angle sum is $\pi + a$ and of which two angles have a sum less than a, so that the third angle would have to exceed π. But this is not possible (why not?—see *122*, 4). Therefore, as always happens in an indirect proof, our entire structure tumbles down, on account of an error in its foundations. We conclude that the angle sum of a triangle can not exceed π.

Theorem LX. If there is any triangle for which the angle sum is equal to two right angles, then this is the case for every triangle.

Proof. (a) If the angle sum of every *right* triangle is equal to π, then it is equal to π for every triangle. The proof of this statement is contained in the paragraph preceding Theorem LVI (see p. 261).

(b) If the angle sum of a right triangle whose right sides are a and b is equal to π, then the same is true for a right triangle whose right sides are pa and qb, where p and q are arbitrary natural num-

bers. For, ABC being a right triangle of angle sum π (see Fig. 54), we obtain a new triangle ABC_1, congruent to ABC, by making AC_1 and BC_1 equal to a and b respectively (3 sides equal; VI). Hence the angle sum of the quadrilateral AC_1BC is 2π. Furthermore each angle of this quadrilateral is a right angle; it is therefore a rectangle (the reader should have no difficulty in proving this; see *122*, 9). Next we construct the triangle BB_1C_1 by producing CB to B_1, so as to make BB_1 equal to a. This triangle is congruent to AC_1B and has therefore also an angle sum equal to π. By continuing in this way, horizontally as well as vertically we can build up a rectangle $CPRQ$, such that $CP = QR = pa$, $CQ = PR = qb$, whose angle sum is 2π. The diagonal PQ divides this rectangle into two right triangles whose right sides are pa and qb, and in each of which the angle sum equals π. For the angle sum can not exceed π in either of them (Theorem LIX); hence, if it were less than π in one of them, the angle sum of the rectangle $CPRQ$ would be less than 2π. Therefore the angle sum of triangle CPQ equals π. Moreover, since every other right triangle whose right sides are pa and qb is congruent to $\triangle CPQ$, it must also have an angle sum equal to π.

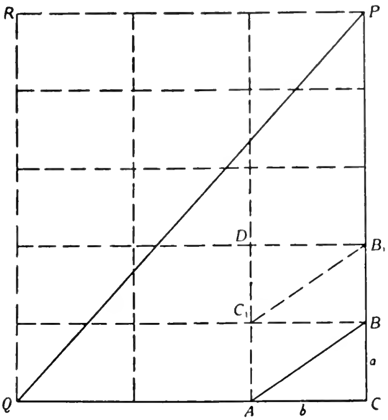


FIG. 54

(c) If there is one right triangle for which the angle sum equals π, then it is equal to π for *every* right triangle. Suppose that the angle sum S of the right triangle PAQ (see Fig. 55) equals π, and that BAC is an arbitrary right triangle. We can then determine natural numbers p and q such that $p \cdot AP > AB$, and $q \cdot AQ > AC$.[1] Let $AP_1 = p \cdot AP$, and $AQ_1 = q \cdot AQ$. We have just seen that the angle sum in $\triangle AP_1Q_1$ must then be equal to π. The line P_1C divides this triangle into the two triangles P_1Q_1C and AP_1C, whose angle sums add up to π + the angle sum of $\triangle AP_1Q_1$, i.e. to 2π.

[1] It should be observed that the axiom of Archimedes is appealed to at this point (compare p. 260, footnote).

Since neither of them can have an angle sum greater than π (Theorem LIX), each of them must have an angle sum equal to π. In particular this is true of $\triangle P_1CA$. But *this* triangle is divided by the line BC into the two triangles P_1BC and ABC. The argument

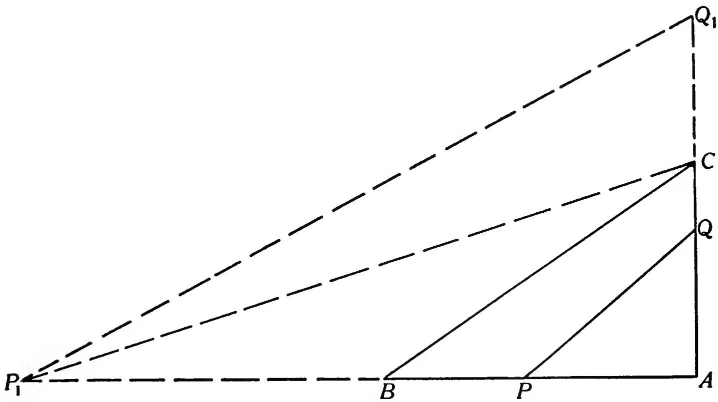


FIG. 55

used just now can therefore be applied a second time; thus we are led to the conclusion that the angle sum of the arbitrary right triangle ABC is equal to π.

(d) If there is any triangle whose angle sum equals π, then there is a right triangle for which this is true (see *122*, 8).

From parts (d), (c) and (a), taken in this order, we obtain the proof of our theorem.

Remark. It is important to observe that Theorems LIX and LX establish the fact that, whether the angle sum of a triangle is less than π or equal to π depends *not* on the particular triangle, but on the "plane" in which the triangle lies as characterized by the earlier postulates. If the angle sum of any one triangle is equal to π, then it is equal to π for every triangle; if, therefore, the angle sum of any one triangle is less than π, then it is less than π for *every* triangle. In other words: *either* the angle sum of every triangle is less than π, *or* the angle sum of every triangle is equal to π. We find a situation entirely analogous to that of Saccheri's Hypotheses I, II and III (compare pp. 258 to 263).

We come next to two theorems which will show that the two alternatives concerning the angle sums of triangles are equivalent

to the alternatives $p(a) = \frac{\pi}{2}$ and $p(a) < \frac{\pi}{2}$, which serve to distinguish between the Euclidean and the Lobatschewskyan geometries.

Theorem LXI. If there are two lines which are perpendicular to the same line and parallel to each other, then the angle sum of every triangle equals π.

Proof. Suppose that AB and PQ are both perpendicular to PA and right parallel to each other (see Fig. 56). We observe in the first place that, although it follows from IV, that if PQ and AB are perpendicular to PA, then those two lines do not intersect, it is by no means certain that they are parallel (compare Definition

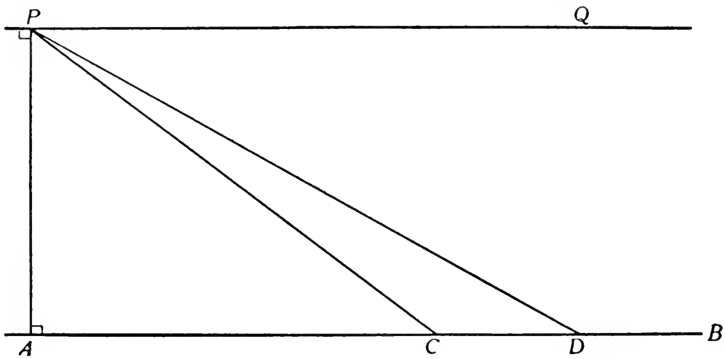


FIG. 56

XXXIX, p. 268); the part of the hypothesis of our theorem which asserts the parallelism of the two lines is therefore not superfluous. Using this assumption, it follows from the same definition that every line through P which lies in the angle QPA will intersect AB. Let PC and PD be two such lines; and let the angle sum of triangles PAC and PCD be denoted by S_1 and S_2 respectively. Then, *either* $S_1 = \pi$ and $S_2 = \pi$, *or* $S_1 = \pi - e_1$ and $S_2 = \pi - e_2$, where $e_1 > 0$ and $e_2 > 0$. If the first alternative holds, our theorem is proved, as a consequence of Theorem LX. Let us suppose then that the second alternative is fulfilled. Then

$$\angle PAC + \angle APC + \angle PCA + \angle PCD + \angle CDP + \angle DPC = 2\pi - (e_1 + e_2); \text{ i.e.}$$

$$\frac{\pi}{2} + \angle APD + \pi + \angle CDP = 2\pi - (e_1 + e_2).$$

From this follows:

$$2\pi - \angle QPD + \angle CDP = 2\pi - (e_1 + e_2), \text{ or}$$

$$\angle QPD - \angle CDP = e_1 + e_2. \tag{11.1}$$

Now, by rotating the line PD about P but keeping it within the angle CPQ, so that it remains a cutting line, the angle QPD can be made as small as may be desired. Consequently, since the difference $\angle QPD - \angle CDP$ appearing on the left side of the equation (11.1) is obviously less than $\angle QPD$, it can certainly be made arbitrarily

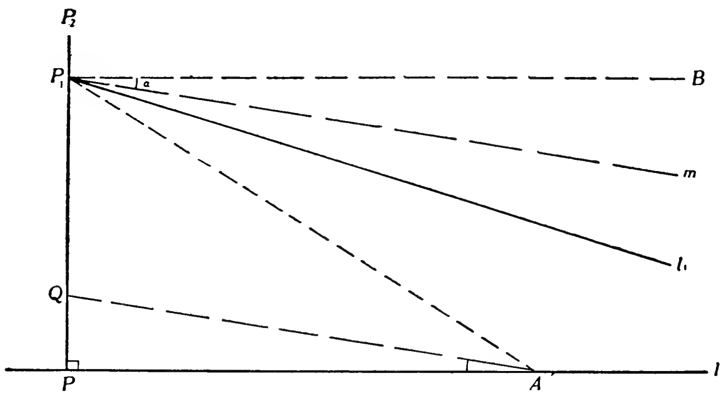


FIG. 57

small; in particular, it can be made less than e_1, which is unaffected by the changing position of PD, since it refers to the fixed triangle APC. On the other hand, the right side of equation (11.1) is always greater than e_1, because $e_2 > 0$. We have thus shown that a contradiction results from the second alternative. Hence the first alternative is the only possible one; the proof of the theorem is therefore complete, since a similar argument can be made if AB and PQ are left parallels.

Theorem LXII. If the angle sum of every triangle is equal to π, then every pair of parallel lines must have a common perpendicular.

Proof. (a) We observe first, without proof (see *122*, 6), that if P_1 and P_2 lie on a line perpendicular to l (see Fig. 57), but so that $P_2P > P_1P$, then the angle of parallelism for P_2 can not exceed the angle of parallelism for P_1.

(b) A second preparatory remark is needed. If the angle sum of a right triangle is π and the sides enclosing the right angle have lengths a and b, then the angles of a right triangle whose right sides have lengths ka and kb, where k is a natural number, are equal to those of the given triangle. The reader will recognize in this statement a simple case of a well-known theorem in the theory of similar triangles. Its proof can be made without the use of the Euclidean parallel postulate by the construction method which was followed in (b) on page 274. The details are left for the reader (see *122*, 11).

(c) Suppose now that l_1 is the right parallel[1] to l through P_1 and that the angle of parallelism $p = \angle PP_1l_1 < \frac{\pi}{2}$, so that the parallels l and l_1 are not both perpendicular to PP_1. We can then draw a non-cutting line m so that $\alpha = \angle PP_1m < \frac{\pi}{2}$ and $\alpha > p$; let $\frac{\pi}{2} - \alpha = \beta$. We connect P_1 with an arbitrary point A on l, to the right of P, so as to form the right triangle PP_1A, in which $\angle PP_1A < p < \alpha$. If now the angle sum of every triangle is equal to π, then $\angle PAP_1 = \frac{\pi}{2} - \angle PP_1A > \frac{\pi}{2} - \alpha = \beta$. Therefore a line AQ through A such that $\angle PAQ = \beta$ will enter the $\triangle PP_1A$ and must therefore cut the line P_1P; let Q be the point of intersection. On the basis of the same hypothesis concerning the angle sum, we conclude then that $\angle PQA = \frac{\pi}{2} - \beta = \alpha > p$. Thus we have through Q a cutting line for l which makes with QP an angle in excess of p.

(d) Now we apply the axiom of Archimedes, which assures the existence of a natural number k such that $k \cdot PQ > PP_1$. We lay off from P on the lines PA and PP_1 (extended) segments PA_1 and PP_2 such that $PA_1 = k \cdot PA$ and $PP_2 = k \cdot PP_1$; the line P_2A_1 completes the triangle PP_2A_1, in which $\angle PP_2A_1 = \angle PQA > p$, on the basis of (b). Thus we obtain through P_2 a cutting line for l which makes with P_2P an angle greater than p. From this it

[1] The argument is not essentially different if l_1 is the left parallel to l through P_1; only one case needs therefore to be considered. In Fig. 57, the angle marked α should be marked β.

would follow that the angle of parallelism for P_2 exceeds p. But since $P_2P > P_1P$, this would contradict (a).

Thus we see that the hypothesis that the angle sum of every triangle is π, is not consistent with the hypothesis that there is a pair of parallels which do not have a common perpendicular. The theorem is therefore proved.

Theorems LXI and LXII are converses of each other. They show, in combination with Theorem LX, that if the angle of parallelism for any point P and any line l is less than $\frac{\pi}{2}$, then no triangle can have an angle sum equal to π; and that, if the angle of parallelism is equal to $\frac{\pi}{2}$, then the angle sum of every triangle is π. We have therefore obtained the following result:

Theorem LXIII. The geometry of a plane in which there exists a triangle whose angle sum is π, is a Euclidean geometry; the geometry of a plane in which there exists a triangle whose angle sum is less than π, is a Lobatschewskyan geometry.

The discussion on the preceding pages should make it clear to the reader that there is nothing mysterious in the geometry of Lobatschewsky; it differs from the familiar Euclidean geometry only in the replacement of the "parallel postulate" by a postulate which asserts the existence of two parallels through every point outside a line. We have gone far enough to see that this replacement has far-reaching consequences. The deductions from the new set of postulates constitute a new geometry and a new trigonometry, which are as consistent in themselves as are the corresponding parts of the classical mathematics. A thoroughly convincing method to establish the inner consistency of the geometry of Gauss, Bolyai and Lobatschewsky on as secure a basis as that of Euclidean geometry has been the outcome of work done by a number of mathematicians during the 19th century; among them must be mentioned Arthur Cayley (1821–1895) and Felix Klein (1849–1925). Of the nature of their work, we shall now try to obtain some idea. This will connect the concepts of Non-Euclidean geometry with common experience and will lead us at the same time to understand the bases for a third type of geometry, with which the name of Bernhard Riemann (1826–1866) is associated. But let us first reflect for a while on the experiences of the present excursion.

122. Meditations in a new world.

1. Prove that through any point P, outside or on the line l, there exists only one line which is perpendicular to l, without making use of Theorem LIX (compare pp. 267 and 272).

2. Show that (a) if $\angle C < \angle B$, then $\angle B_1AC < \angle B_1$; (b) if $\angle C > \angle B$, then $\angle B_1AC > \angle B_1$ (see Fig. 53, p. 272).

3. Prove that for any line l and any point P outside l, the two angles of parallelism are equal (compare p. 269).

4. Prove that no angle of a triangle can exceed two right angles, without using Theorem LIX (compare p. 273).

5. Prove that if three angles of a quadrilateral are right angles, then the fourth angle can not be obtuse.

6. Prove that if $P_1P_2P \perp l$ and $P_1P > P_2P$, then the angle of parallelism for P_1 can not exceed the angle of parallelism for P_2 (compare p. 277 (a)).

7. Prove the corollary of Theorem LVIII (see p. 272).

8. Prove that if there is any triangle whose angle sum equals π, then there is a right triangle for which this is true (compare p. 275 (d)).

9. Prove that all the angles of the quadrilateral AC_1BC in Fig. 54 are right angles.

10. Prove that the quadrilateral ACB_1D in Fig. 54 is a rectangle.

11. Prove that if the right sides of a right triangle have lengths a and b and if its angle sum equals π, then the angles of a right triangle whose right sides are ka and kb are equal to those of the given triangle, for any natural number k (compare p. 278 (b)).

123. Geometry on an orange. It was shown in Theorem LXIII that the angle sum of a triangle can serve as a distinguishing mark between the two types of geometry with which we have now become somewhat familiar. It was also pointed out that the angle sum will be less than π or equal to π according to the character of the "plane" in which the triangle lies (compare pp. 262 and 275). What then is a "plane"?

In so far as the "plane" has been used in our geometrical deductions, it has no physical properties by which it can be characterized and recognized. It is true that we have used the surface of some of these pages as the medium in which to represent the illustrative diagrams. But for this purpose we have not drawn upon their physical properties, not any more on their flatness than upon their whiteness or glossiness. The only properties of the plane which we have used are those inherent in the first four postulates which have served as the basis of our work. Consequently a "plane" is

anything which has related to it elements which we may call "points" and "lines" (without presupposing for these elements any of the physical characteristics usually associated with these names) in such a way that postulates I–IV hold. Let us recall what these postulates are by reading the footnote on page 254 and let us then try to determine different "planes" for which they are satisfied.

The simplest example of this kind is that already referred to, viz. the surface of a sheet of paper, with one important modification, however. When we "draw a line" and "mark off a point" by means of pencil and straight edge, we obtain the elements which are ordinarily called lines and points.[1] Are the first four postulates satisfied by these lines and points in this plane? The first one manifestly holds, for we can draw one of our straight lines from any point to any point. Moreover the definitions of straight angle and right angle have evident interpretations from which the equality of all right angles follows, so that the fourth postulate also holds. There is some difficulty with the second and the third, however. For we can not produce a finite straight line in our plane beyond the edges of the plane. Neither can we describe a circle in the plane whose radius exceeds half the width of the plane. To remedy these defects we may think of the sheet of paper which is serving as our plane, as being extended indefinitely beyond its edges. With this extension of the first concept of the plane, we obtain indeed an example in which the first four postulates are satisfied. What can we say about the geometry in this plane? Is it Euclidean or Non-Euclidean? If we apply the criterion of the angle sum of a triangle a difficulty arises, because actual determination of this angle sum may be troublesome. It is as conceivable that this angle sum will turn out to be less than π, as that it will be π. There is therefore no reason why a Non-Euclidean geometry might not arise in this example. It is of importance to remember that, to make postulates I to IV significant, we had to extend the plane beyond all bounds; this could only be done by *thinking of it* as of infinite size.

Let us see what happens to postulates I to IV if we take the surface of a sphere as our "plane." It is clear enough what we are to

[1] It should be understood throughout this paragraph that whenever we speak of line or point we mean the pencil marks that are commonly designated by these words, even though we do not use the quotation marks.

take for the points. What about the lines? To obtain a suitable definition for them, we have to recall a few simple facts from solid geometry. If a plane is made to cut a sphere the points common to the plane and to the surface of the sphere (the curve of intersection) always form a circle (compare Fig. 58). The radius of this circle depends upon the distance of the sphere's center from the intersecting plane. The nearer the plane is to the center of the sphere, the greater its radius. It is at its maximum when the plane passes through the center of the sphere; in this case the radius of the circle of intersection is equal to the radius of the sphere. Therefore the circles obtained by the intersection of a sphere with planes through its center have radii equal to the radius of the sphere. With evident justification these circles on the surface of the sphere are called the *great circles*. If we take now two arbitrary points P and Q on the surface of the sphere, such that the line PQ does not pass through the center, there is exactly one plane which passes through P and Q, and through the center of the sphere. In accordance with what has just been seen, we can conclude therefore that *through two arbitrary points on the surface of a sphere there passes exactly one great circle,* provided these points are not diametrically opposite points. It is therefore reasonable, if we want our postulates I–IV to hold, to define the great circles on the sphere as its "straight lines." We have now the necessary elements for a geometry on the surface of a sphere. Moreover postulate I is satisfied except for two diametrically opposite points. What about postulate II? It clearly does not hold for if we produce one of our spherical straight lines to a length equal to $2\pi \times$ the radius of the sphere, we come back to the starting point. Hence in the geometry on the surface of a sphere of radius R, it is not possible to continue a straight line if its length be $2\pi R$, not at least in the ordinary sense of these words (compare *110* (b),

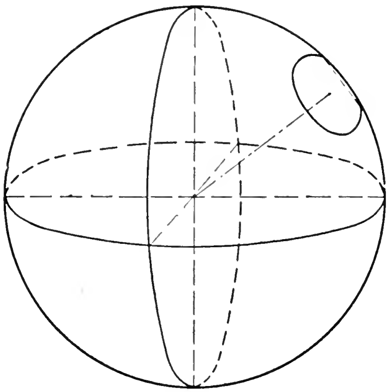


FIG. 58

p. 243). Since therefore postulates I to IV are not all valid, neither the conclusions of Lobatschewsky nor those of Euclidean geometry need hold; in particular, it should not cause surprise if Theorem LIX (see p. 272) were to fail. This does actually happen; in the "plane" geometry which we have established on the sphere, the angle sum of a triangle exceeds π. A proof is found in every book on solid geometry, which can readily be understood by any one who will take the trouble to study the elements of that subject. Such a person will moreover discover the remarkable fact that the "area" of a spherical triangle depends upon the amount by which its angle sum exceeds π. The reader can easily make these facts plausible to himself if he will but take an orange and recognize that on its surface he can cut three grooves in the position of great circles, which are two by two perpendicular to each other; thus he will obtain a triangle of which each of the angles is a right angle, so that the angle sum equals $\frac{3\pi}{2}$. If he now cuts additional grooves on the surface of the orange through one of the vertices of his triangle (compare Fig. 59), still at right angles to the opposite side, he will obtain a succession of triangles whose angle sums decrease from $\frac{3\pi}{2}$ towards π and whose areas decrease from one eighth of the area of the surface of the orange to 0.

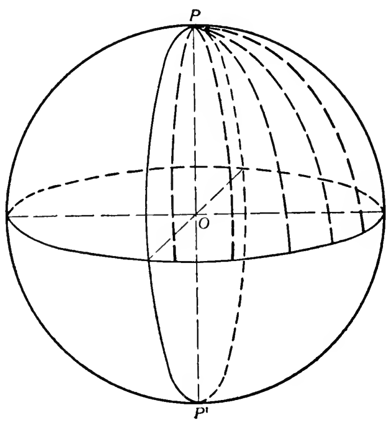


FIG. 59

Thus far we have not exhibited a "plane" on which the angle sum of the triangles is actually less than π. The surface of the sphere is a "plane" in which the angle sum of the triangles is greater than π. This "plane" has been introduced because it is easily visualized and because its study paves the way for the consideration of some further facts which are necessary for the construction of a "plane" of the other type. Our concern with the

sphere will assist us in gaining some understanding of the additional concepts they involve.

124. Outlook on further possibilities. The different circles which have been drawn through the point P on the surface of the sphere in Fig. 59 (the grooves on the orange) all turn the same way, viz. downward from the point P. Indeed all the planes which pass through the vertical diameter POP' of the sphere cut its surface in circles which bend in this way. Suppose that, instead of the sphere, we take an ellipsoid (we can think of this as a sphere flattened in one or two directions; see Fig. 60). If through a point P on the surface of this solid we draw a line at right angles to the surface,[1] and then pass planes through this line, the planes cut the surface of the ellipsoid in curves which in general are not circles, but which nevertheless all bend at P in the same sense.

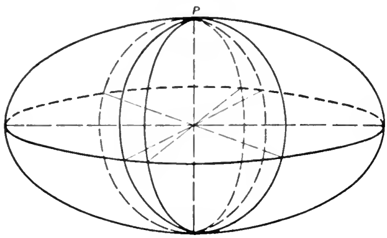


FIG. 60

Let us consider on the other hand a surface which is shaped like a saddle (see Fig. 61) and let us carry out the same constructions on it. It should be clear to the reader (particularly if he supplements the illustration by making a model out of a malleable substance like clay), that some of the curves that are so obtained turn downwards, while others turn upwards. We distinguish between the two types of surfaces illustrated in Figs. 60 and 61 by calling one of the former type a *surface of positive curvature*, and one of the latter type a *surface of negative curvature*.[2] It is not difficult to imagine surfaces

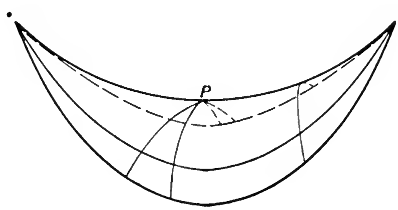


FIG. 61

[1] An exact discussion of the angle between a line and a surface can not be given at this point (compare p. 335); we appeal to the reader's geometric intuition for an understanding of what is meant.

[2] If the surface is of such character that the curves traced on it do not bend at all, i.e. if these curves are straight lines, we call it a *surface of zero curvature*. Our ordinary plane is a surface of this kind.

on which there are some points at and near which the shape of the surface is like that of a piece of a sphere or an ellipsoid, while it also contains other points at and near which the surface is shaped like a saddle surface, and again, points at and near which the surface is like a plane. In other words, we can think of surfaces which have positive curvature at some points, negative curvature at other points, and zero curvature at still other points. But such surfaces we can leave out of consideration for our present purpose. We are concerned with surfaces which have everywhere positive curvature, with surfaces which have negative curvature at every point, and with such as have zero curvature at every point.

Just as we have been able to introduce "straight lines" on the sphere, so we can also characterize curves on surfaces of less simple shape, which fulfill, at least in part, the rôle of straight lines; they are called the geodesics of the surface. Having done this, we can establish a geometry on surfaces of constantly positive curvature, such as spheres and ellipsoids for instance, and a geometry for surfaces of constantly negative curvature of which the saddle surface furnishes an example. The detailed study of these geometries requires a larger knowledge of technical mathematics than has so far been acquired. A number of the concepts involved will be taken up in subsequent chapters; indeed the present discussion is unavoidably vague because these concepts are not at our disposal. Therefore we have to content ourselves now with a mere statement of the following facts:

1. The geometry on a surface of constantly negative curvature is such that the angle sum of a triangle is always less than π; it is a Lobatschewskyan geometry, provided this curvature is of constant magnitude.

2. The geometry on a surface of constantly positive curvature is such that the angle sum of a triangle is always greater than π.

3. In the geometry on a surface of zero curvature the angle sum of a triangle is equal to π; it is a Euclidean geometry.

This is one of the important results of the work referred to at the end of *121*. To Riemann is due the discovery of a geometry in which, as in the geometry on the surface of a sphere, the angle sum of a triangle exceeds π.

To obtain these new "plane geometries" we have made use of surfaces which are situated in the 3-dimensional space of our daily experience. This should not blind us to the fact that the surfaces

are in themselves 2-dimensional. For organisms living on and confined to such surfaces there would be no way of determining whether they were on a surface of positive curvature or on one of negative curvature. At least they could not make such a distinction in the way in which it has been made here, because this involves 3-dimensional considerations. From the vantage point which we occupy we could supply them with the means of settling their dilemma. We would tell them to find out whether the angle sum of their triangles is less than π, or greater than π, or possibly equal to π.

The Non-Euclidean geometries which we have now interpreted as geometries on surfaces of positive or of negative curvature in our ordinary 3-dimensional space are thus seen to be just as real and just as self-consistent as our plane geometry (at least as our solid geometry), which is now recognized as the geometry on a surface of zero curvature.[1] They represent an enlargement of our experience, which had heretofore been limited, as far as the study of 2-dimensional geometry is concerned, to surfaces of one type only, viz. to planes. As we free ourselves from this limitation and take into account surfaces of positive and negative curvature, we recognize the possibility of new "plane" geometries which realize in a concrete way theories which had been conceived of initially as purely logical structures. It is important to recall that this liberation had its starting point in breaking away from the rather vague physical image which we had of a "plane" and in considering it as any entity in which we can introduce "points" and "lines" in such a way that at least some of the first four postulates of Euclid are satisfied. It is in this way that we arrived at a solution of the age-old problem to which Euclid's parallel postulate gave rise.

Furthermore the reader will recognize that there is no a priori reason why we should limit ourselves to surfaces of constantly positive or constantly negative curvature. By widening our field of study so as to include surfaces of still different character there is a possibility of creating many more types of geometry. Along this road modern mathematics has attained an understanding of Euclidean geometry not as the one and only geometry but as one in a vast collection of geometries.

[1] In the study of surfaces, there are considered several kinds of curvature. What we have here called "curvature" of a surface is technically known as *total curvature* or *Gaussian curvature.*

125. What kind of space do we live in? Before proceeding with the consideration of the problems to which our initiation into Non-Euclidean geometry has led us we must pause for a few moments to reflect a little further on another interesting aspect of this study.

We are dwellers not merely on the surface of the earth, but also inhabitants of an ever increasing portion of the space which surrounds the earth (and of the inside as well — think of miners, divers, swimmers and submarines). That is to say we are three-dimensional beings, and not two-dimensional creatures. It is due to this fact probably that the human race became able after many centuries of labor to conceive of different kinds of two-dimensional worlds. But there is another side to the story. While our three-dimensionality is of great advantage for the understanding of two-dimensional worlds, it involves a serious handicap in the consideration of worlds of more than three dimensions, and even in the attempt to gain insight into the three-dimensional universe. In this respect we are handicapped in the same way as the two-dimensional creatures introduced to the reader on page 286.[1] Having recognized the possibility of different kinds of two-dimensional worlds, each with its own geometry and with its own supply of furniture which depends upon its geometry, it is natural that we should inquire whether there may not also exist various kinds of three-dimensional worlds, each with its own geometry and connected property. Above all we should then like to know the character of the three-dimensional world to which the space of our experience leads us. Indeed it is this question with which physicists, astronomers and mathematicians have been very much concerned in recent times and to which important contributions have been made. The effort has been to devise some scheme which, like the determination of the angle sum of a triangle in the "plane," would make it possible to decide what sort of space it is we live in. The development of mathematics has brought within the reach of man the conception of different types of spaces of three (and of more) dimensions, a study of their properties and means of testing to which type of space our world corresponds. The experiments to which these studies have led are by no means complete, and it is not at all certain whether an absolutely definitive answer can be found. Many scientists are engaged in work in this field. Among them the name of Albert Einstein (1879–) has become widely

[1] Pleasant and instructive reading on this subject is *Flatland*, by E. A. Abbott.

known; of the many important contributors to this work we mention also Hermann Weyl (1885–), Hermann Minkowski (1864–1909), Willem de Sitter (1872–1934) and Arthur Eddington (1882–). Any account of their work is out of the question here. It will be sufficient for our purpose if the reader has gained some idea of the nature of the problem with which it is concerned.

CHAPTER XII

TO THE HEADWATERS OF A GREAT RIVER

Every one who understands the subject will agree that even the basis on which the scientific explanation of nature rests, is intelligible only to those who have learned at least the elements of the differential and integral calculus, as well as of analytical geometry. — Felix Klein, *Jahresbericht der Deutschen Mathematiker Vereinigung*, 1902, p. 131.

126. Planning the itinerary. To obtain examples of Non-Euclidean planes we have made use of surfaces of positive and of negative curvature. This was done on the basis of only a very vague notion of what is meant by the curvature of a surface. While it is beyond our purpose to enter into a detailed study of this question, it is desirable to see at least what is involved in such an inquiry. This is the aim of the present expedition. We have to obtain a wide view over a large region in order to trace the sources of the stream which we have seen in its lower course — various methods of travel will have to be used.

Let us begin a little farther back and consider first some simpler matters. "Reculer pour mieux sauter" is a useful plan in intellectual matters as well as in track athletics. On pages 284–286 we introduced the concepts of positive and negative curvature of a surface by a consideration of the plane curves obtained as the intersection of the surface with planes through a line at right angles to the surface. The first step in our study must therefore relate to the curvature of a curve at a point.

127. What is a curve? The question which has just been raised is one that has occupied many people these last fifty years. Doubtless everyone can give some intuitive answer. It is the sort of thing which we "know perfectly well" but which we find great difficulty in expressing. Felix Klein has written "Everyone knows what a curve is until he has studied enough mathematics to become confused through the countless number of possible exceptions."[1] Our experience of moving objects, of the flight of a bird, of the

[1] See Moritz, *Memorabilia Mathematica*, p. 329.

motion of a ball through the air, of a fish through the water, of a shooting star through the night sky, of a child's hand trying to grasp an object, provides us with sufficient background to give intuitive significance to the word "curve." An exact definition of this concept which is useful for mathematical purposes has been elaborated in recent times; we may come back to it later on but we shall not try to develop it now. For the present it will suffice if we can agree that essential elements involved in the concept are change of position and perhaps of direction, and that by change of position direction is generated.

If we limit ourselves to curves in a plane, position can be specified by means of a system of coördinates (compare p. 21). How about direction?

The general concept *curve* which we are considering would include that of a straight line in the sense of Euclidean geometry; it is the simplest example of a "curve"[1]; and for it the meaning of direction is quite clear. If two straight lines enclose an angle we say that they have different directions; if they are parallel we describe them as having the same direction. This suggests that when we want to deal with the direction of a single line we have to compare its position with that of another line. Thus we could specify the direction of a line l by means of the angle which it makes with some fixed line in its own plane. For some purposes this may be quite sufficient. Our aim is served better if we introduce another quantitative measure for the direction of a straight line; this will at the same time give us a good start on a more systematic discussion of curvature.

128. The meaning of slope. We have already had to do with directed lines, when we studied vectors in Chapter V; at that time we made use of a system of coördinates. We return now once more to this system of coördinates in a plane, with the notations and terminology which were described and used in Chapter II; it is convenient to add to our vocabulary the words *abscissa* and *ordinate* which are synonymous with x-coördinate and y-coördinate respectively. On the line l we take two points, as A and B (see Fig. 62). The directed segment AB has the components AB_1 and B_1B (compare p. 98). The ratio of these two components is a number. This number is sufficient to characterize the direc-

[1] It is clear that "curve" in this technical sense is more inclusive than in its common usage, in which it is contrasted with straight line.

tion of the line, for: (1) it is independent of the particular directed segment of the line which, like AB, may be chosen to determine this ratio, and (2) it is the same for two lines if and only if these lines are parallel.[1] It is easy to substantiate these two statements. The first follows from the fact that, for any

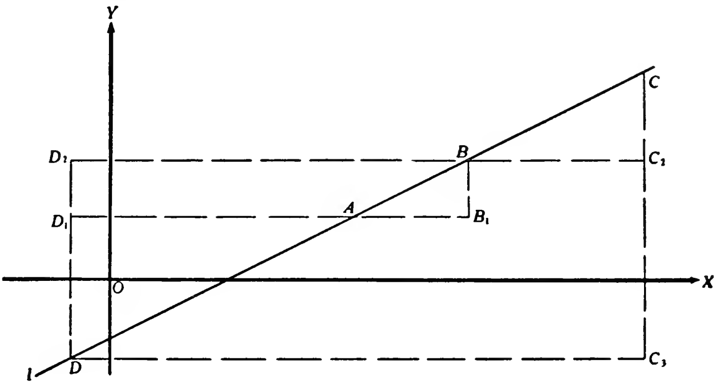


FIG. 62

two directed segments of l, the right triangles formed by the segments and their components are similar; so, e.g., AB_1B, DC_3C, BD_2D are all similar.[2] The second is proved by noticing (compare Fig. 63) that the parallelism of the lines l and m is equivalent to the equality of the angles at A and at P, hence to the similarity of the right triangles AB_1B and PQ_1Q and finally to the equality of the component-ratios $\dfrac{B_1B}{AB_1}$ and $\dfrac{Q_1Q}{PQ_1}$.

This component-ratio is taken as a measure of the direction of the line l and is called its *slope*.

Definition XLI. The slope of a line with respect to a system of rectangular coördinates is the ratio of the Y-component to the X-component of any directed segment of the line.

The following considerations are of importance for our further work.

1. If a line passes through the origin it will enter *either* quad-

[1] Throughout our work from here on, we shall use Euclidean geometry unless the contrary is specifically stated.

[2] Care must be taken to observe the sense of the segments and to give the correct algebraic sign to the components; e.g. the components AB_1, C_2C are positive, but the components AD_1 and D_2D are negative.

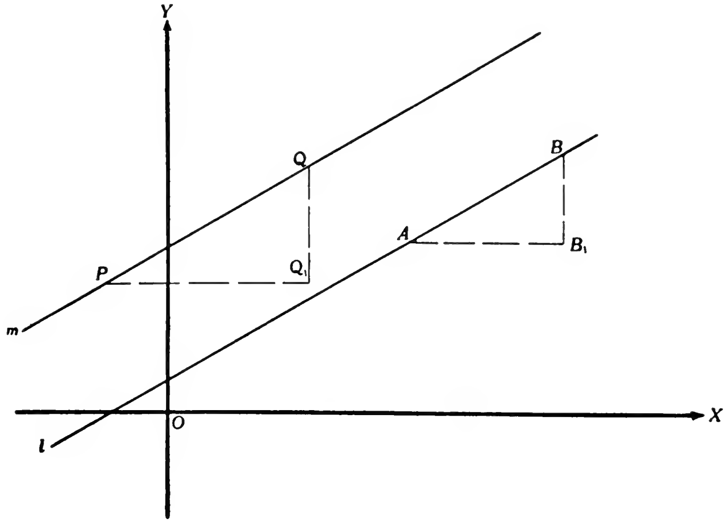


FIG. 63

rants I and III, or quadrants II and IV.[1] In the former case the two components of any of its directed segments are of like sign (both positive or both negative), so that its slope is positive. In the latter case, the two components of any of its directed segments are opposite in sign, and its slope is therefore negative. If a line l does not pass through the origin, its slope will be positive or negative according as the line through O and parallel to l runs into quadrants I and III, or into quadrants II and IV.

2. This discussion does not include the X- and Y-axes nor the lines parallel to them. For the X-axis and lines parallel to it the Y-component of any segment is 0; hence their slope is zero. For the Y-axis and lines parallel to it, the X-component of every segment is zero, so that the slope of such lines is non-existent. (Compare p. 39, where division by 0 is discussed.)

3. The slope of a line l with respect to a system of rectangular coördinates is evidently dependent upon the angles the line makes with the coördinate axes. Of these two angles we consider in particular the angle α from the X-axis to l (see Fig. 64). If we draw a line l_1 through O parallel to l, and determine the slope of

[1] Compare p. 22.

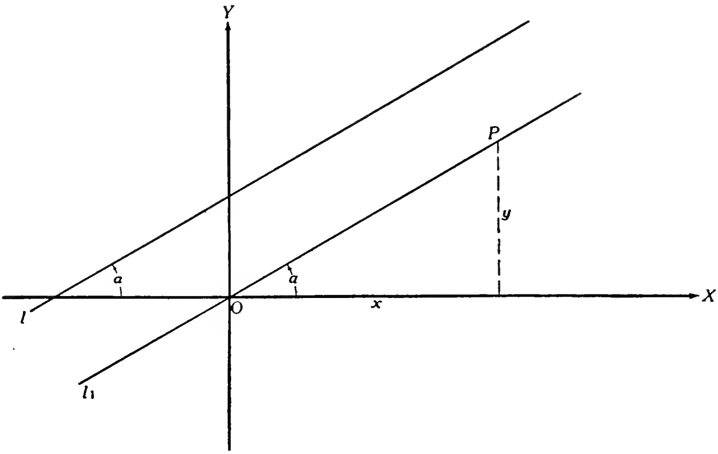


FIG. 64

the latter by means of a directed segment from the origin O to a point P, whose coördinates are x and y, we obtain for the slope the quotient $\frac{y}{x}$ of these coördinates. The reader will remember that the sine and cosine of the angle, as defined in Chapter V (compare p. 98), would turn out to be equal to $\frac{y}{OP}$ and $\frac{x}{OP}$ respectively. If we divide sin α by cos α we obtain

$$\frac{\frac{y}{OP}}{\frac{x}{OP}},$$

which reduces to $\frac{y}{x}$, that is to the ratio which determines the slope of l_1. We define now a new trigonometric function of the angle α as follows:

Definition XLII. The tangent of an angle α is equal to the quotient of its sine by its cosine; it is denoted by the abbreviation tan α.

On the basis of this definition we can say that the slope of the line l_1 through the origin is equal to the tangent of the angle from

the X-axis to l_1; but since the slope of l is equal to that of l_1, and since the angle from the X-axis to l is equal to that from the X-axis to l_1, we conclude that the slope of any line l is equal to the tangent of the angle from the X-axis to l.

We summarize this discussion as follows:

Theorem LXIV. The Y-axis and lines parallel to the Y-axis have no slope; the X-axis and lines parallel to it have slope 0. Lines through the origin which enter into the 1st and 3rd quadrants, and lines parallel to them, have positive slope; lines through

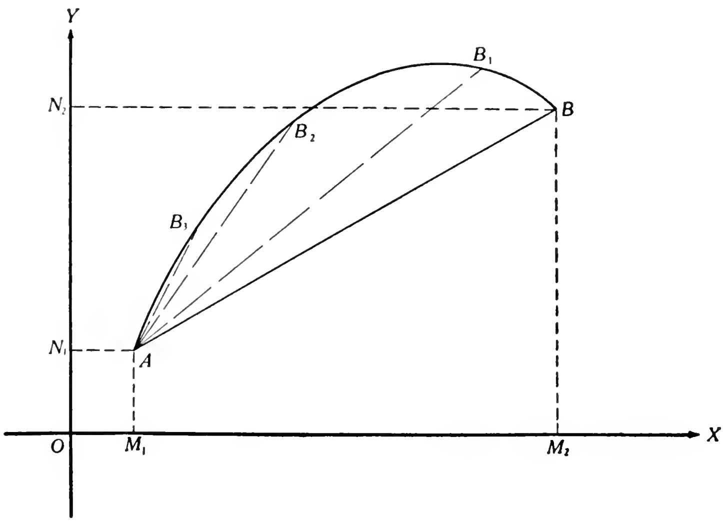


FIG. 65

the origin which enter into the 2nd and 4th quadrants, and lines parallel to them, have negative slope. Conversely, lines of zero slope are parallel to or coincident with the X-axis, lines of positive slope pass through the origin and enter the 1st and 3rd quadrants or are parallel to such lines; lines of negative slope pass through the origin and enter the 2nd and 4th quadrants or are parallel to such lines. Lines for which the slope is non-existent are coincident with or parallel to the Y-axis. The slope of any line is equal to the tangent of the angle from the X-axis to the line.

129. Direction of a curve. We shall now carry over the concept of direction from a straight line to a curved line. In Fig. 65,

we have drawn an arbitrary curved line. What are we to understand by the direction of this curve at the point A, and how shall we measure it? Doubtless the reader has an intuitive notion of what he means by the direction of the curve at A; let us try to analyze this intuitive notion and to give it explicit formulation. If we connect A with another point B on the curve by means of a straight line, the direction of this straight line AB gives something approximating the desired direction. If we take the point B_1 on the curve nearer to A than B is, the direction of the line AB_1 is felt to be a better approximation than the direction of the line AB to what we intuitively conceive to be the direction of the curve A. Now it may happen, and with the curves that we ordinarily have to deal with it always will happen, that, as we take points B_1, B_2, B_3 etc. closer and closer to A, the slopes of the corresponding lines AB_1, AB_2, AB_3 etc. will differ less and less from some fixed number. For instance, these slopes may be $\frac{1}{2}$, $\frac{3}{4}$, $\frac{7}{8}$, $\frac{15}{16}$ and so forth. More precisely it may, and usually will, happen that, by taking a point B_n near enough to A, we shall obtain for the line AB_n, and for all lines joining A to points closer to A than B_n is, slopes which differ from some fixed number a by less than any preassigned positive amount no matter how small. In this case we shall say that the slope of the curve at the point A is equal to a.

After this preliminary discussion we can proceed to an explicit formulation. As always, in order to obtain exact statements, we must begin with some definitions.

Definition XLIII. If a variable number v changes in such a way that the numerical value of its difference from a fixed number a *becomes and remains* less than any arbitrarily preassigned positive number e, a is called the limit of v.

As entirely equivalent to the statement "a is the limit of v," we shall use phrases like "v approaches a," or "v tends towards a"; the notations most frequently used are $\text{Lim } v = a$ and $v \longrightarrow a$.

By the numerical value of a positive or negative number is meant, as is well known, the unsigned number obtained by omission of the $+$ or the $-$ sign. The notation $|a|$ is used to designate the numerical value of a; thus, $|-\frac{2}{3}| = |+\frac{2}{3}| = \frac{2}{3}$ etc. (compare p. 88).

Examples of this sort of variability come readily to mind. If v runs through the sequence of values $\frac{1}{2}$, $\frac{3}{4}$, $\frac{7}{8}$, $\frac{15}{16}$ etc., its limit is 1; for the numerical values of the differences $1 - \frac{1}{2}$, $1 - \frac{3}{4}$, $1 - \frac{7}{8}$,

$1 - \frac{15}{16}$ etc. become and remain less than any e.[1] If v runs through the sequence $1\frac{1}{2}$, $1\frac{1}{4}$, $1\frac{1}{8}$, $1\frac{1}{16}$ etc. it also has the limit 1; in this case the differences between 1 and the successive values of v are all negative, and the *numerical* value of the differences again becomes and remains less than any e. The same is true if v runs through the following sequence: $\frac{1}{2}$, $1\frac{1}{2}$, $\frac{3}{4}$, $1\frac{1}{4}$, $\frac{7}{8}$, $1\frac{1}{8}$, $\frac{15}{16}$, $1\frac{1}{16}$ etc.; or through $\frac{1}{2}$, 1, $\frac{3}{4}$, 1, $\frac{7}{8}$, 1, $\frac{15}{16}$, 1, etc.; or through $\frac{2}{3}$, 1, $1\frac{1}{3}$, $\frac{8}{9}$, 1, $1\frac{1}{9}$, $\frac{26}{27}$, 1, $1\frac{1}{27}$ etc.[2] On the other hand if v runs through the sequence 1, 2, $1\frac{1}{4}$, $1\frac{3}{4}$, $1\frac{1}{8}$, $1\frac{7}{8}$, $1\frac{1}{16}$, $1\frac{15}{16}$ etc. it does not approach any limit; neither if its values are 1, 2, 3, 4, 5 etc. We have had important examples of variables approaching a limit in Chapter IV (compare Theorems VI, p. 53; VI*a*, p. 54; VII, VIII, p. 56). In terms of the notation and terminology which we have introduced, Definition XLIII can be restated as follows:

Definition XLIIIa. "The variable v approaches the limit a" is equivalent to the statement that $|v - a|$ becomes and remains less than e.

If the concept limit of a variable has been thoroughly absorbed, we are ready for the next step. It frequently happens that two variables, let us call them x and y, are so related that as $x \to a$, y will tend to b. Suppose e.g. that $y = x + 2$; then as $x \to 1$, it should be clear that $y \to 3$. Whenever this is the case, we record the phenomenon in the following symbolic way:

$$\lim_{x \to a} y = b; \quad \text{or thus:} \quad y \to b \text{ as } x \to a.$$

Of especial interest are variables which tend to 0; on account of their frequent occurrence, a special name is given to them.

Definition XLIV. An *infinitesimal* is a non-constant variable which tends to zero.

The reader should have no difficulty in supplying examples of

[1] To avoid long statements we shall understand that e, or sometimes ϵ, stands for an "arbitrarily preassigned positive number, no matter how small." When there is danger of confusion with the Napierian base e, we shall use ϵ.

[2] To forestall difficulties arising from erroneous conceptions unfortunately current in many elementary textbooks which introduce the concept of limit, it should be observed that it is a matter of complete indifference whether the "variable reaches the limit"; that is to say, it does not matter in the least whether in the course of its history the variable v does or does not take the value a. By carefully examining the examples given in the text, the reader will become aware of this fact. An extreme case of a variable v which tends to a would be that in which it takes the values a, a, a etc., i.e. the case in which v is constantly equal to a. For also in that case the conditions of Definition XLIII are satisfied; and this is the one and only thing that is required.

infinitesimals. It is to be carefully noted that a constant no matter how small can not be an infinitesimal. We have here another instance of the deviation from the common meaning when a word is used in a technical sense (compare p. 290, footnote).

Observe that the phrase "x is an infinitesimal" means that $|x|$ becomes and remains less than e.

We are now prepared to make a precise formulation of the meaning of the direction of a curve at a point.

Definition XLV. If A is a point on a curve and if the slope of the straight line AB tends to a definite limit whenever the distance from B to A is made infinitesimal by variation of B along the curve, this limit is called the *slope of the curve at* A.

Definition XLVI. If A is a point on a curve, the straight line through A whose slope is equal to the slope of the curve at A is called the *tangent line* of the curve at A.

If we combine Definitions XLV and XLVI with Definition XLII, we obtain the following equivalent formulation for the slope of a curve at a point: The slope of a curve at the point A is the tangent of the angle from the X-axis to the tangent line of the curve at A.

Our inquiry as to the direction of a curve at a point A has therefore received the following answer: The direction of a curve at a point A is the direction of the tangent line of the curve at that point; it is measured by the slope of the tangent line.

We have completed the first stage of our expedition: much remains to be done. But it is time for some recreation.

130. Practice in preparation for the next advance.

1. Draw lines through the origin whose slopes are $\frac{2}{3}$; $-\frac{3}{2}$; $\frac{5}{4}$; $-\frac{4}{5}$.

2. Draw lines through the point $P(-2, 1)$ whose slopes are 1; $-\frac{3}{4}$; -2; $\frac{1}{2}$.

3. Determine the slopes of the lines which join

(a) $P_1(-1, 4)$ and $Q_1(3, -1)$;
(b) $P_2(-2, -1)$ and $Q_2(3, 2)$;
(c) $P_3(-4, 3)$ and $Q_3(4, -3)$;
(d) $P_4(2, 3)$ and $Q_4(0, -2)$.

4. Show that if the slope of a line is $+1$, the line makes an angle of $45°$ with the X-axis: if the slope is -1, the line makes an angle of $135°$ with the X-axis.

5. What is the slope of the X-axis and of lines parallel to it?

6. Show that the slope of the line which passes through the origin and through the point $P(a, b)$ is equal to $\frac{b}{a}\cdot$

7. Show that $\tan 0° = 0$, $\tan 45° = 1$, $\tan 135° = -1$, $\tan 180° = 0$, $\tan 225° = 1$, $\tan 315° = -1$, $\tan 360° = 0$. What can you say concerning $\tan 90°$ and $\tan 270°$?

8. Make use of the data given on pages 99 and 100 to determine $\tan 30°$, $\tan 120°$, $\tan 150°$, $\tan 200°$, $\tan 300°$.

9. Show that if θ is an angle in I or III, then $\tan \theta$ is positive, and if θ is an angle in II or IV, then $\tan \theta$ is negative (compare p. 97); and conversely.

10. Prove that if l and m are lines through the origin which are mutually perpendicular, and if the slope of l is equal to $\frac{b}{a}$, then the slope of m is $-\frac{a}{b}\cdot$

11. Prove: if l and m are lines through the origin and if the slope of l is equal to $\frac{b}{a}$ and that of m is equal to $-\frac{a}{b}$, then l and m are mutually perpendicular.

12. Prove that if two lines are mutually perpendicular the product of their slopes is equal to -1; and, conversely, if the product of the slopes of two lines is equal to -1, the lines are mutually perpendicular.

13. Show that if $y = x^2$ and $x \to 2$, then $y \to 4$.

14. Show that if $P(a, b)$ is a point whose coördinates satisfy the relation $y = x^2$, then the slope of the line OP is equal to a.

15. Show that if $P(a, b)$ and $Q(c, d)$ are two points whose coördinates satisfy the relation $y = x^2$, then the slope of the line PQ is equal to $c + a$.

16. Show that if P and Q are any two points whose coördinates satisfy the relation $y = 2x + 3$, then the slope of PQ is always equal to 2.

17. Prove that if P and Q are any two points whose coördinates satisfy the relation $y = ax + b$ (in which a and b are real numbers), then the slope of the line PQ is always equal to a.

18. Show that if $P(a, b)$ and $Q(c, d)$ are two points whose coördinates satisfy the relation $xy = 1$, then the slope of PQ is always equal to $-\frac{1}{ac}\cdot$

19. Determine the slope of the line PQ if the coördinates of P and of Q satisfy the relation $y = x^3$.

20. Determine the slope of the line PQ if the coördinates of P and Q satisfy the relation $y = 3x^2 - 4x$.

131. Restrictions of liberty. If a point P is allowed to wander freely over the plane, its abscissa and its ordinate can vary at liberty; each of them can become equal to any real number,

independently of the other. This fact is expressed in the statement that, for an unrestricted point P, the x- and y-coördinates are *independent variables* over the class C of real numbers (compare pp. 63, 65); also in the phrase "the plane is a set of points with *two degrees of freedom.*"

However, if the point P is restricted to a curve, the freedom of its coördinates is considerably curtailed. For the curve drawn in Fig. 65, the abscissa of P may vary over the set of real numbers represented by the points between M_1 and M_2; but, once the abscissa of P has been chosen, there is no liberty left for its ordinate — there is only one possible value for it and that is dependent on the choice of the abscissa. On the other hand the ordinate of P may vary over the set of real numbers represented by the points between N_1 and N_2; but, when it has been fixed, the abscissa is restricted to one, or at most two, possibilities, which depend upon

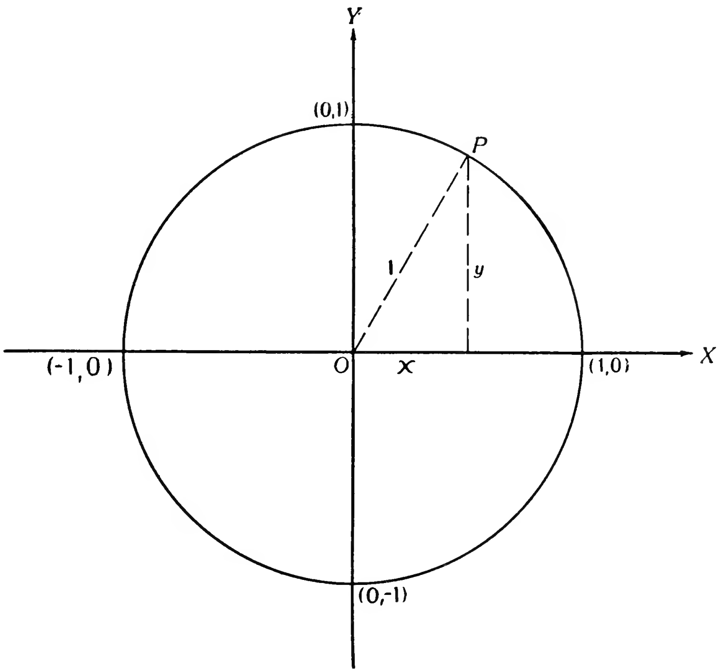


FIG. 66

the choice of the ordinate. Let us consider a few other examples of curves, in order to discover the state of liberty of the coördinates of a point P whose wanderings they control.

If the curve is a circle of radius 1 about the origin, and the x-coördinate of P is left as much freedom as is consistent with circumstances, this coördinate may vary over the set of real numbers from -1 to $+1$, including the ends. But to each value of x between -1 and $+1$ there correspond only two values of the y-coördinate; to the values -1 and $+1$ of x, only one value of y, viz. 0. Similarly, the y-coördinate has freedom to vary from -1 to $+1$; but then the x-coördinate is restricted to one or two values. There is therefore liberty, although not complete liberty, for either one of the coördinates; the other one depends upon it (see Fig. 66).

If the curve is the one represented in Fig. 67 (which is understood to continue indefinitely to the right and to the left, upwards

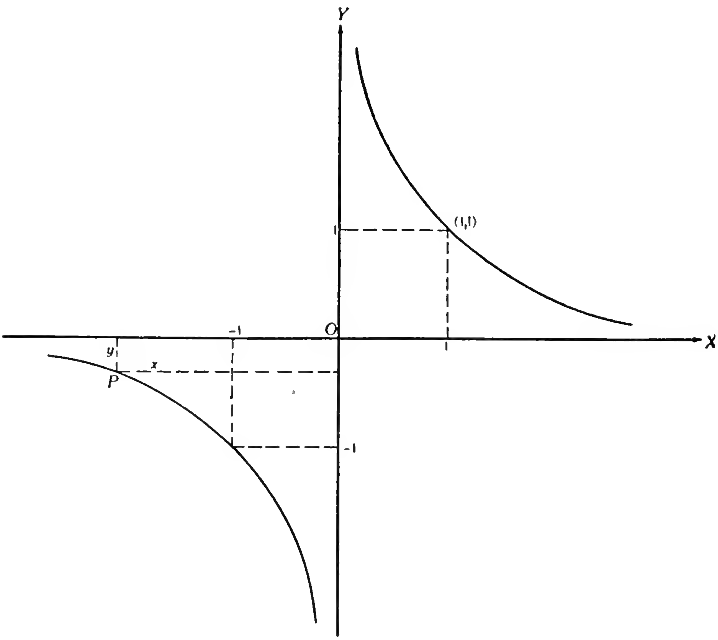


FIG. 67

and downwards, approaching the coördinate axes, but having no point on either axis), the x-coördinate is free to roam over the set C of real numbers, with the sole exception of the value 0; but with each value of x there is associated but a single value of y. And if the y-coördinate is left independent, it can take any real value, (except 0), and with each value of y there is associated a single value of x. Again we see that, if P is restricted to this curve, either coördinate is independent, while the other depends upon it.

For the curve suggested in Fig. 68 (it is intended to wind around the Y-axis indefinitely in both directions, like an idealized bean vine), the x-coördinate is free to vary between -1 and $+1$; and although to each such value of x there corresponds a denumerable infinitude of values of y, the variability of y, once the x has been chosen, is definitely restricted. On the other hand if y has the

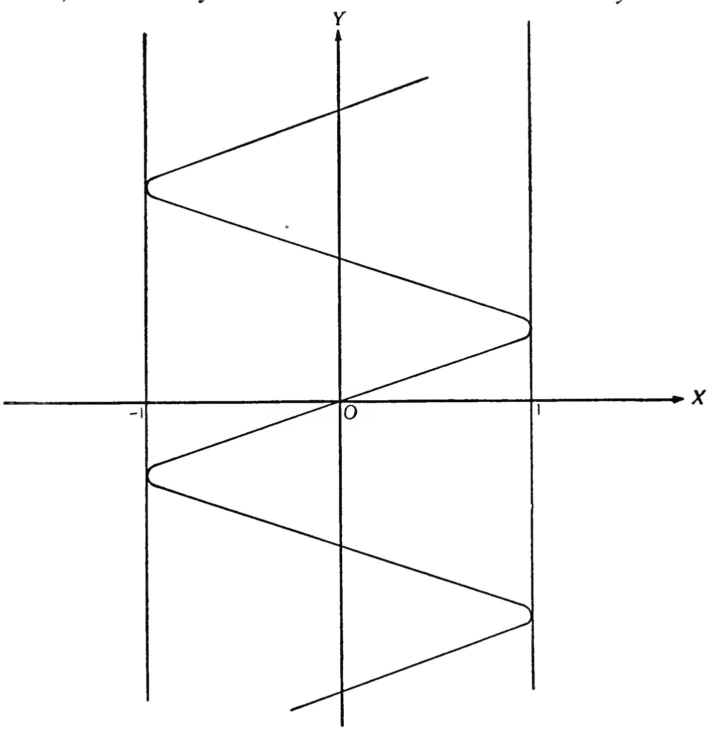


FIG. 68

first choice it is free to take any real value; and with every such value of y there is associated a single value of x.

These examples of curves are clearly not sufficient to justify a general conclusion, especially since we are still handicapped by a lack of definiteness in our concept of a curve. They suggest, however, a possible way in which this concept may be specified. Following this suggestion, we shall introduce a definition of curve, which has by no means the generality of the conceptions to which allusion was made at the beginning of *127* (see p. 290), but which will carry us a long way in our present undertaking.

Definition XLVIIa. A plane curve is a set of points P of the plane obtained when one of the coördinates of P is allowed to vary over the set C of real numbers (or over a subset of C consisting of the real numbers between two specified real numbers a and b, including both, one or neither end) and when to each value of this coördinate corresponds one, a finite number or a denumerable infinitude of values of the other coördinate.

132. Functional relations resulting from restrictions. This definition is certainly sufficiently general to include the examples of curves which have just been given. We shall restrict our study to a less general class of curves, viz. to those in which to each value of the coördinate of P which is left free there corresponds *exactly one* value of the other coördinate. We shall call such curves *simple plane curves*, according to the following definition.

Definition XLVIIb. A simple plane curve is a set of points P of the plane obtained by allowing one of the coördinates of P to vary over the set C of real numbers (or over a subset of C, as in Definition XLVIIa), and by associating with each value of this coördinate *one* value of the other coördinate.

In accordance with these definitions the curves represented in Figs. 67 and 68 are both simple plane curves, but the circle drawn in Fig. 66 is not a simple plane curve. We do obtain simple plane curves when we take the upper (or the lower) semi-circle, also by taking the right half (or the left half) of this circle.

For these simple plane curves we can now obtain an analytic representation. For every point P of such a curve we have a pair of numbers, viz. the x- and y-coördinates of P. As P varies along the curve, both x and y go through a set of values; x and y are then variables. One of them can range over the whole or a part of the set C of real numbers; to each value of that one corresponds

one value of the other. To make matters definite let us suppose that the x is left free; we shall call it the *independent variable.* With each value of x there is associated a single value of y; it is called the *dependent variable.* Thus for each curve there exists a relation between the two variables x and y, such that to each value of the independent variable x there corresponds one value of the dependent variable y. A relation of this kind is called a *functional relation.* The connection between x and y which has here been discussed at such length is described in technical language, by the statement: "the dependent variable y is a function of the independent variable x"; and it is written symbolically in the form $y = f(x)$, which is read "y is a function of x." We summarize the discussion as follows:

Definition XLVIII. When the two variables x and y are so related that to each value of x chosen from the whole or a part of the set of real numbers C there corresponds a single real value of y, we say that y is a function of x; we represent the relation in the form $y = f(x)$.

For every simple plane curve there is therefore a relation of the form $y = f(x)$ or a relation of the form $x = f(y)$. Frequently, but not always, it is possible to express the functional relation between y and x which corresponds to a given curve by means of simple equations. It is readily seen that for the curve (1) represented in Fig. 69, i.e. for the straight line through the origin which makes an angle of 45° with the X-axis, the corresponding relation takes the simple form $y = x$; this is called the equation of the line. For the curve (2), which is parallel to (1) but cuts the Y-axis at the point (0, 2), the equation is $y = x + 2$; for the curve (3), it is $y = x - 3$. By reference to our starting point, we find that the upper half of the circle in Fig. 66 has the equation $y = \sqrt{1 - x^2}$, where the symbol $\sqrt{P}$ denotes, as always, the positive square root of P; the lower half has the corresponding equation $y = -\sqrt{1 - x^2}$. The right half of the same circle leads to $x = \sqrt{1 - y^2}$; the left half to $x = -\sqrt{1 - y^2}$.

The systematic study of the relations between x and y which correspond to given curves belongs to the science of *Analytical Geometry.* It can not be our purpose to undertake such a study; that would be a departure from our general aim. Those readers who feel attracted to this subject can find many books devoted to

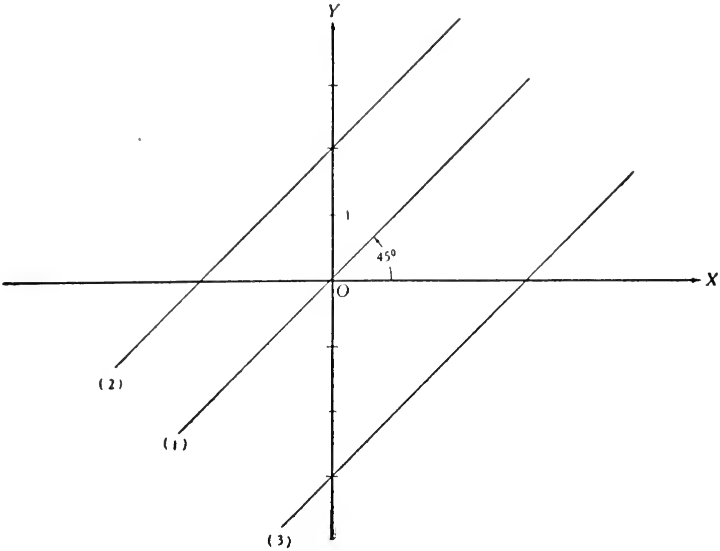


Fig. 69

it. Among the books available in English we mention the following: Fine and Thompson, *Coördinate Geometry;* Salmon, *Conic Sections*. There is another fundamental problem of analytical geometry that we must mention, viz. the inverse of the one just stated (remember Jacobi!). If a functional relation between the variables x and y is given by means of algebraic operations (or operations of a different nature), what is the curve corresponding to it? For instance if x and y are so related that $y = x^2$, what curve corresponds to this relation? That to any relation, either of the form $y = f(x)$ or of the form $x = f(y)$, there corresponds a curve, should be clear to any one who has understood definitions XLVII*b* and XLVIII.

The variables x and y, in terms of which the definition and the examples of functional relations have been stated, may represent measurements of a great variety of scalar magnitudes. Simple illustrations of functional relations from various fields of science readily suggest themselves: the relations between the area and the radius of a circle, between the pressure of a constant volume of gas and its temperature, between the temperature of a metal wire and the

intensity of an electrical current passing through it, between barometric pressure and altitude above sea-level, between the intensity of cosmic radiation and geographical latitude, etc., etc. Examples of a more complicated character have arisen in the modern analysis of economic phenomena. It should be clear that the study of functional relations between variables, which is undertaken in mathematics without regard to the possible interpretations of these variables, brings the subject into intimate connection with many fields of science. In each case the functional relation gives expression to a "law" in a scientific domain; the corresponding curve exhibits the law graphically. There are many instances in which a curve has been obtained as a result of scientific observations. The determination of the corresponding functional relation leads to the formulation of the scientific law which expresses the rules governing the observations and which makes possible the prediction of new phenomena.

The correspondence between curves (geometric entities) and functional relations (analytical entities) lies at the basis of analytical geometry. The fundamental idea from which this subject has been developed is usually attributed to the French philosopher and mathematician René Descartes (1596–1650).[1]

133. Some technical equipment essential for the explorer. Before coming to the central problem of this chapter, viz. the calculation of the slope of a curve, it will be well to have a look at a few simple facts from analytical geometry. This will also give us some idea of the methods used in this subject.

Theorem LXV. To any straight line in the plane there corresponds an equation in x and y in which these variables occur at most to the first degree.

Proof. (a) If l is a line through the origin (but not the Y-axis), which makes the angle θ with the X-axis and P is an arbitrary point on l (see Fig. 70), then $\dfrac{y}{OP} = \sin\,\theta$, $\dfrac{x}{OP} = \cos\,\theta$; consequently $\dfrac{y}{x} = \tan\,\theta$ (compare Definition XXII, p. 98, and Definition XLII, p. 293). Hence, for any point P on the line and for no other point, the coördinates satisfy the relation

$$y = x\tan\,\theta. \tag{12.1}$$

[1] Compare Cajori, *op. cit.*, p. 185, and corresponding places in other books on the history of mathematics.

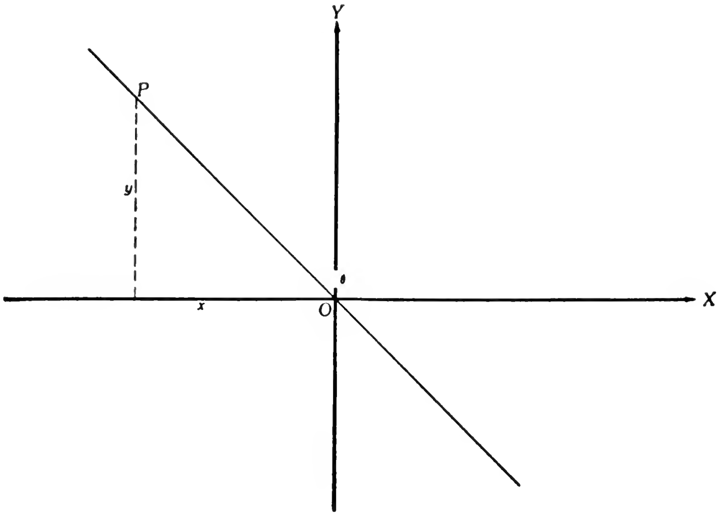


FIG. 70

(b) If l is an arbitrary line (not parallel to the Y-axis), which makes an angle θ with the X-axis, we draw a line l_1 through the origin and parallel to l (see Fig. 71). Let B be the point in which l meets the Y-axis and let $OB = b$ (notice that OB is a directed line; in Fig. 71, b is positive). If P is an arbitrary point on l of coördinates (x, y), and PQ is drawn parallel to the Y-axis, this line will meet l_1 in a point P_1, whose coördinates are (x, y_1) and such that $P_1P = b$. We find that $y = QP = QP_1 + P_1P = y_1 + b$ if and only if P is on the given line l; but it was found in (a) that $y_1 = x \tan \theta$. We conclude that the equation of the line l is

$$y = x \tan \theta + b. \tag{12.2}$$

We observe now that in equations (12.1) and (12.2), the variables x and y occur to the first degree; moreover (12.2) reduces to (12.1) when b is replaced by 0. If the line l is parallel to the X-axis and at a distance b from it, $\theta = 0$; hence $\tan \theta = 0$ (compare *130*, 7); the relation (12.2) becomes $y = b$, in which x and y occur to the first degree at most. For the X-axis itself we find $y = 0$.

(c) It remains to consider lines l parallel to the Y-axis and the

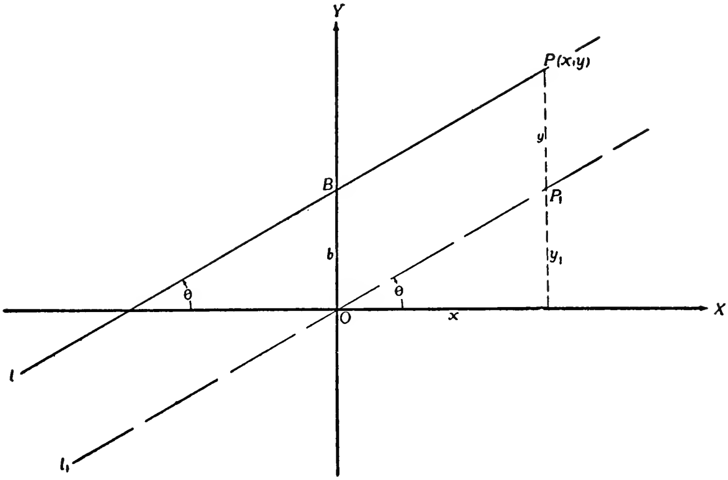


FIG. 71

Y-axis itself. If the distance of l from the Y-axis is c, every point on l has its x-coördinate equal to c, while this is not true for points which are not on l. The required equation for such a line is then $x = c$; if $c = 0$, the line l coincides with the Y-axis for which we have then $x = 0$. These relations have again the required form.

Thus the theorem has been completely proved.

Theorem LXVI. To any equation in x and y with real coefficients in which these variables occur to at most the first degree, but not both to the degree zero, there corresponds a straight line.

Proof. The most general equation of the type mentioned in the theorem is $px + qy + r = 0$, in which p, q and r are real numbers, not both p and q being zero. If $q \neq 0$, the equation can be reduced to the form

$$(12.3) \qquad y = -\frac{px}{q} - \frac{r}{q}.$$

This has exactly the form (12.2), provided we can determine an angle α such that $\tan \alpha = -\frac{p}{q}$. We shall assume for the present that this determination is indeed possible. It then follows from our earlier work that to the equation $px + qy + r = 0$ in which

$q \neq 0$ corresponds the straight line which cuts the Y-axis at a point B, such that $OB = -\dfrac{r}{q}$ and for which the angle α with the X-axis is determined by the relation $\tan \alpha = -\dfrac{p}{q}$.

If $q = 0$, the equation (12.3) has no meaning, but the original equation becomes $px + r = 0$, in which $p \neq 0$. It reduces to the form $x = -\dfrac{r}{p}$; to this equation corresponds a line parallel to the Y-axis (compare (c), above).

Hence the proof is complete.

Remark. The two preceding theorems have established a one-to-one correspondence between the set of straight lines in the plane and the set of equations in x and y of the form $px + qy + r = 0$, in which p and q are not both equal to zero. To every straight line in the plane corresponds one such equation; to every such equation corresponds one line. The equation which corresponds to a given line is the *equation of the line;* the line is said to be the *locus of the equation.* Thus $y - x = 0$ is the equation of the line through the 1st and 3rd quadrants which bisects the angle between the X- and Y-axes; the Y-axis is the locus of the equation $x = 0$, etc. It should now be clear why equations of the type $px + qy + r = 0$ which we have been considering are called *linear equations in x and y.*

From the equation of a line we can moreover determine the slope of the line; this conclusion follows at once from the results which have just been obtained, in connection with Theorem LXIV. We state it as follows:

Theorem LXVII. The locus of the equation $px + qy + r = 0$ has no slope if $q = 0$; in all other cases its slope is equal to $-\dfrac{p}{q}$.

To determine the angle which the line makes with the X-axis, we have to fill up the gap left in the proof of Theorem LXVI, by showing that, whenever p and q are two real numbers, there always exists an angle α such that $\tan \alpha = -\dfrac{p}{q}$. To do this, we have but to join to the origin the point, whose coördinates are $(q, -p)$, see Fig. 72, or any other point whose x- and y-coördinates are proportional to q and $-p$ (e.g. the point $(-q, p)$). Thus we obtain

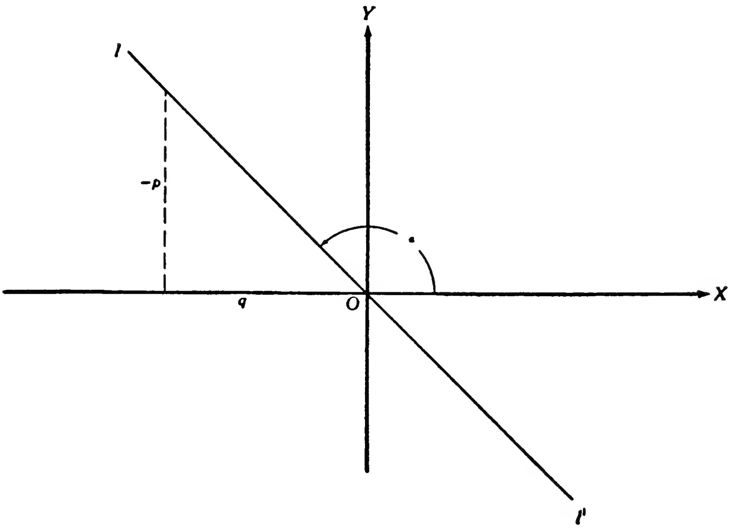


FIG. 72

an undirected line through the origin. The reader should have no difficulty in showing that for the angle XOl or the angle XOl' the tangent is equal to $-\dfrac{p}{q}$.[1] It is verified incidentally that the slope of a line is not affected by a change of direction on the line.

In the study of the correspondence established between the straight lines of the plane and the linear equations in x and y we have an illustration of the procedure of analytical geometry. It involves general methods to determine the equation of a curve, given in the plane, and to discover the geometric properties of the locus of a given equation in x and y. For a linear equation $px + qy + r = 0$, the locus is always a straight line. If $q \neq 0$ the slope of the line is $-\dfrac{p}{q}$, and the distance it cuts off from the Y-axis is $-\dfrac{r}{q}$; if $q = 0$, the line is parallel to the Y-axis. When it is put in the form (12.3), the equation $px + qy + r = 0$ is seen

[1] Use should be made of Definition XLII (see p. 293); for the actual determination of the angle α the discussions and tables on pp. 93–99 will be found useful; see also *130*, 6.

to be a special case of the general relation $y = f(x)$, which corresponds to a simple plane curve (compare p. 302).

A few further theorems are needed before we can proceed.

Theorem LXVIII. The slope of the line which joins the points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ is equal to $\frac{y_2 - y_1}{x_2 - x_1}$.[1]

Proof. It follows from what is given that the X- and Y-components of the segment P_1P_2 are equal to $x_2 - x_1$ and $y_2 - y_1$ respectively. Consequently we obtain in virtue of Definition XLI (see p. 291) for the slope of the line the number $\frac{y_2 - y_1}{x_2 - x_1}$, as was to be proved.

Theorem LXIX. The equation of the line which joins the points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ is $(x_2 - x_1)(y - y_1) = (y_2 - y_1)(x - x_1)$.

Proof. (a) If the line P_1P_2 is not parallel to the Y-axis and if $P(x, y)$ is an arbitrary point on it, the slope of the line can be expressed in terms of the X- and Y-components of the segments P_1P and P_1P_2; the equation as given results from the statement of the equality of these two expressions. It is satisfied if and only if P lies on the line P_1P_2. Hence this equation is the equation of the line P_1P_2.

(b) If the line P_1P_2 is parallel to the Y-axis, we have $x_1 = x_2$ and $y_1 \neq y_2$; the equation stated in the theorem reduces then to the form $x - x_1 = 0$, whose locus is indeed the line through P_1 parallel to the Y-axis (compare the last paragraph in the proof of Theorem LXVI).

Corollary. The equation of the line through the point $P_1(x_1, y_1)$ with slope m is $y - y_1 = m(x - x_1)$.

We have acquired some new technique — the minimum amount needed for the expedition; some familiarity with it must be secured before we can proceed with safety.

134. Going in for training.

1. A circle is constructed with center at the origin of coördinates and with a radius equal to 4. The ordinates of all points on the upper half of this circle are then reduced to half their size. Determine the relation between x and y which corresponds to the curve that is obtained in this way.

[1] This formula presupposes that $x_2 \neq x_1$. If $x_2 = x_1$, the line in question is parallel to the Y-axis and has therefore no slope.

2. The abscissas of all points on the left half of the circle described in 1 are reduced to $\frac{3}{4}$ of their size. Determine the relation between x and y for the new curve.

3. In the circle described in 1, diameters are drawn which divide the circle in 12 equal sectors, beginning with the X-axis. Determine the equations of the lines on which these diameters lie.

4. Construct the loci of the following equations:

(a) $3x + 2y - 5 = 0$; (c) $2x - 3y + 5 = 0$;
(b) $3x - 5 = 0$; (d) $3y - 5 = 0$.

5. Decide, without constructing the loci, whether the lines represented by the following equations pass through quadrants I and III, or through quadrants II and IV:

(a) $5x + 2y = 0$; (c) $y = \frac{2x}{3}$;
(b) $3x - 4y = 0$; (d) $4x + 5y = 0$.

6. For each of the following equations determine the position of lines through the origin which are parallel to their loci:

(a) $2x - 7y + 11 = 0$; (d) $5x + 2y - 7 = 0$;
(b) $x = \frac{4y}{5} - \frac{3}{5}$; (e) $3y - \frac{x}{5} + 2 = 0$;
(c) $y = -\frac{2x}{3} + \frac{1}{2}$; (f) $\frac{2x}{5} + \frac{7y}{5} - 3 = 0$.

7. Draw the lines which are the loci of the following equations:

(a) $3x - 2y + 1 = 0$; (d) $\frac{x}{3} + \frac{y}{4} = 1$;
(b) $y = \frac{x}{2} - \frac{5}{2}$; (e) $4x + 2y - 11 = 0$;
(c) $5x + 3y - 7 = 0$; (f) $x = \frac{3y}{4} + \frac{1}{4}$.

8. The points $A(-2, 5)$, $B(3, 1)$, $C(0, -4)$ and $D(-5, 0)$ are connected in the order A, B, C, D. Show that the quadrilateral formed by these points is a parallelogram.

9. Determine the equations of the diagonals of the quadrilateral specified in 8.

10. Determine the equations of the lines which join

(a) $P_1(-1, 4)$ and $Q_1(3, -1)$;
(b) $P_2(-2, -1)$ and $Q_2(3, 2)$;
(c) $P_3(-4, 3)$ and $Q_3(4, -3)$;
(d) $P_4(2, 3)$ and $Q_4(0, -2)$.

11. Show that if a line has positive slope then the ordinates of its points increase (decrease) when the abscissas increase (decrease); and that if the slope of a line is negative, then the y-coördinates and x-coördinates of its points vary in inverse senses.

12. Show that the lines $3x + 4y - 2 = 0$, $3x + 4y + 18 = 0$, $4x - 3y + 5 = 0$ and $4x - 3y - 15 = 0$ enclose a square.

135. A rapid march to the source. The preparations are now complete — we can start.

(a) Let us begin by considering the special case of the point $A(1, 1)$ on the curve given by the equation $y = x^2$. (See Fig. 73.)

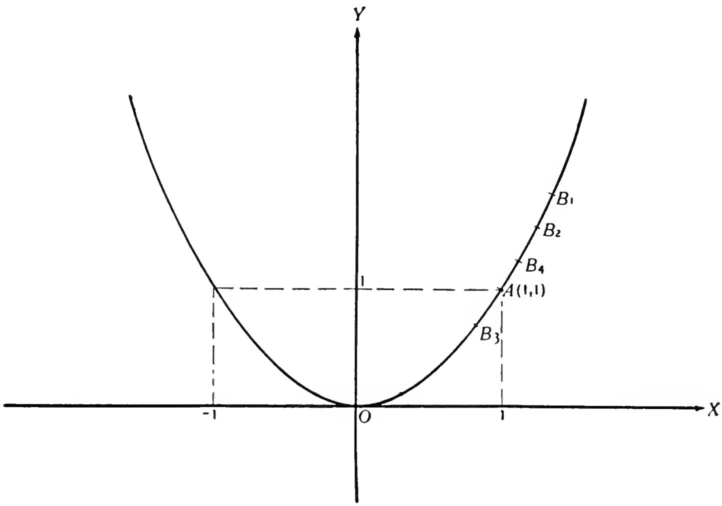


FIG. 73

In accordance with Definition XLV (see p. 297), we have to consider the slopes of the straight lines AB_1, AB_2, AB_3, . . . obtained by joining A successively to *points on the curve* which tend toward A (compare p. 295). If we denote the coördinates of the point B_n by $(1 + h_n, 1 + k_n)$, then we can express (1) the fact that B_n lies on the curve by the statement that

$$(12.4) \qquad 1 + k_n = (1 + h_n)^2, \quad n = 1, 2, 3, \ldots;$$

and (2) the condition that B_n is a variable point which tends toward A by the requirement that $h_n \to 0$, i.e. by the statement that h_n is an infinitesimal. Observe that, for various values of n, the points

B_n can lie on either side of the point A; for the requirement that $h_n \to 0$, means that $|h_n|$ becomes and remains less than e, so that h_n may be negative as well as positive. We have then to determine the slope of the line from $A(1, 1)$ to $B_n(1 + h_n, 1 + k_n)$ and to inquire what becomes of it as $h_n \to 0$. In accordance with Theorem LXVIII (see p. 310), the slope of the line AB_n is equal to $\frac{k_n}{h_n}$; this quotient can be expressed in terms of h_n alone by means of the condition (12.4). This equation takes the simple form

$$k_n = 2h_n + h_n^2;$$

from this we conclude[1] that

$$\frac{k_n}{h_n} = 2 + h_n.$$

It is now easy to determine what becomes of the slope $\frac{k_n}{h_n}$ of the line AB as $h_n \to 0$. In fact, if we have understood the discussion in *129* we should see at once that $\frac{k_n}{h_n} \to 2$ as $h_n \to 0$. Thus we have found that the slope of the curve represented by the equation $y = x^2$ at the point $A(1, 1)$ is equal to 2.

To obtain the equation of the line tangent to the curve $y = x^2$ at the point $A(1, 1)$ we make use of the Corollary of Theorem LXIX (see p. 310). By means of it we find that the equation of the tangent line is $y - 1 = 2(x - 1)$.

(b) From this very special case we advance to one somewhat more general. Let us consider the point $A(a, a^3)$ on the curve whose equation is $y = x^3$ (see Fig. 74). Let the points $B_1, B_2, \ldots B_n, \ldots$ be represented by $(a + h_1, a^3 + k_1), (a + h_2, a^3 + k_2), \ldots (a + h_n, a^3 + k_n), \ldots$. The fact that these points lie on the curve $y = x^3$ finds expression in the equation

$$(12.5) \qquad a^3 + k_n = (a + h_n)^3, \quad n = 1, 2, \ldots;$$

to insure that they tend toward A, we have but to require that $h_n \to 0$. We proceed now exactly as before. The slope of the line AB_n is given in the first place by $\frac{k_n}{h_n}$. This number is ex-

[1] From the equation $k_n = 2h_n + h_n^2$ we deduce also the fact that $k_n \to 0$ as $h_n \to 0$, i.e. that k_n is an infinitesimal when h_n is an infinitesimal.

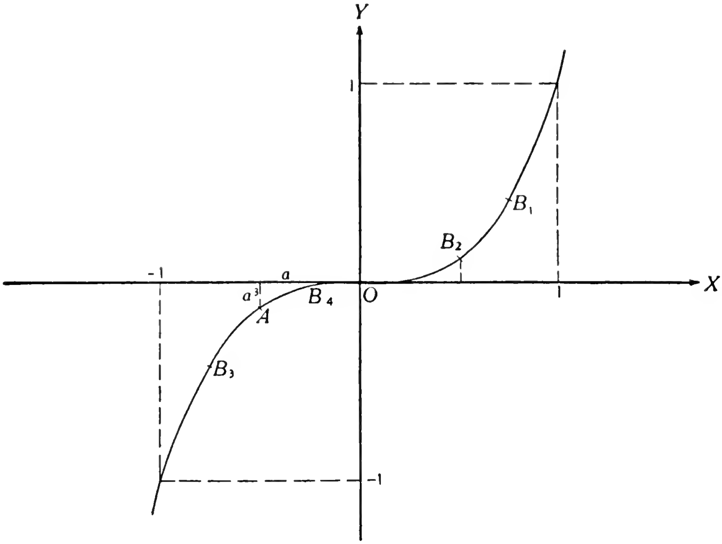


FIG. 74

pressed in terms of h_n alone by use of the equation (12.5); from it we find that

$$k_n = 3a^2h_n + 3ah_n^2 + h_n^3,$$

and hence that $$\frac{k_n}{h_n} = 3a^2 + 3ah_n + h_n^2.$$

Now we apply the method discussed in *129*, and we conclude[1] that as $h_n \to 0$, $\frac{k_n}{h_n} \to 3a^2$. Therefore the slope of the curve $y = x^3$ at the point (a, a^3) is equal to $3a^2$. By selecting arbitrary values for a, we can readily calculate the slope of the curve at any of its points. For the equation of the tangent line to the curve at the point (a, a^3), we find $y - a^3 = 3a^2(x - a)$, or $3a^2x - y - 2a^3 = 0$.

[1] To obtain this result the remarks made in *129* are not sufficient. We have to use in addition to them the following theorems: 1. The product of two infinitesimals is an infinitesimal; 2. The product of an infinitesimal by a constant is an infinitesimal; 3. The sum of two infinitesimals is an infinitesimal. The reader should have no difficulty in seeing how these theorems are used in the present instance. Proofs can be found in a few of the books alluded to on p. 318; statements without proof in all others. It will be worth while to attempt independent proofs (compare *138*, 20, 23, 25). Observe also that, as in (a), $k_n \to 0$ as $h_n \to 0$.

Two general properties of the curve $y = x^3$ can be deduced from the fact that its slope is given by $3a^2$.

1. The slope of this curve is never negative. The origin is the only point at which it is equal to zero; at every other point it is positive. It follows that at the origin the curve is tangent to the X-axis, and that at every other point the curve rises as it moves to the right. (Compare Theorem LXIV, p. 294; and *134*, 11.)

2. The slope of the curve gets larger as the point moves farther away from the origin, either to the right or to the left; i.e. the curve gets steeper in either direction away from the origin.

(c) We should now be ready for the general case of a simple curve (compare *132*), represented by an equation of the form

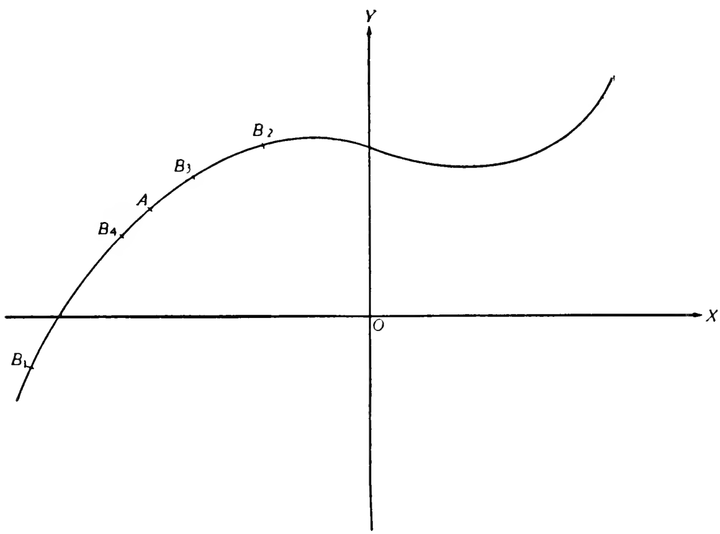


FIG. 75

$y = f(x)$, see Fig. 75. Let the point A be represented by the coordinates $[a, f(a)]$ and the points B_1, B_2, . . ., etc. by $[a + h_1, f(a) + k_1]$, $[a + h_2, f(a) + k_2]$, . . ., etc. respectively. The fact that the points B_n lie on the curve whose equation is $y = f(x)$, finds expression in the relation

$$f(a) + k_n = f(a + h_n), \quad n = 1, 2, \ldots; \tag{12.6}$$

and we can indicate that these points tend towards A, by requiring that h_n be an infinitesimal.[1] The slope of the line AB_n is found again, by use of Theorem LXVIII, to be equal to $\frac{k_n}{h_n}$; and from (12.6) it follows that this is equal to

$$\frac{f(a + h_n) - f(a)}{h_n}.$$

If we remember now the definition of the slope of a curve at a point (see Definition XLV, p. 297), we can justify the following result:

Theorem LXX. If the fraction

$$\frac{f(a + h_n) - f(a)}{h_n}$$

tends towards a definite limit as h_n tends to zero, then the slope at the point $[a, f(a)]$ of the curve represented by the equation $y = f(x)$ is equal to that limit; if this fraction does not tend towards a limit as h_n tends to zero, then the curve does not possess a slope at that point.

This theorem furnishes an explicit analytical method for calculating the slope of a curve at a point; in so far as we can apply it, we are able to solve the problem mentioned at the beginning of *133*. The problem has lost its geometrical character and has received an analytical formulation; the procedure illustrates once more the methods of analytical geometry (compare the end of *132*).

If we think of the relation between the variables x and y, expressed by the equation $y = f(x)$, without regard to a geometric interpretation, the limit of the fraction

$$\frac{f(a + h_n) - f(a)}{h_n}$$

as h_n tends to zero is called the instantaneous rate of change of the variable y with respect to the variable x at the value a of x. The reason and justification for this terminology can readily be appreciated as follows. The slope of a line can obviously be interpreted as the ratio of the change in the variable represented by y to the

[1] In this general case it is not evident, as it was in (a) and (b) that k_n is an infinitesimal, when h_n is an infinitesimal. We shall make the *assumption* that this is the case for all the functions $f(x)$ which occur in our work; compare *140*.

corresponding change in the variable represented by x, as we pass from one point on the line to another point on the line, i.e. as the *rate of change* of y with respect to x. If the two variables are related by a linear equation, this rate of change is constant. In the general case represented in Fig. 75, the slopes of the lines AB_1, AB_2, . . ., represent the *average* rate of change of y with respect to x for changes in x represented by the abscissa differences of these pairs of points. If this average rate tends to a limit as the abscissa difference tends to 0, this limit is the instantaneous rate of change.

For this purely analytical concept the term *derivative* is used.

Definition XLIX. If the fraction

$$\frac{f(a + h_n) - f(a)}{h_n}$$

tends to a limit, as h_n tends to zero, this limit is called the *derivative with respect to* x of the function $f(x)$ at $x = a$. It is frequently denoted by the symbol $D_x f(x)$, at $x = a$; or by the symbol $f'(a)$.

The derivative of a function $f(x)$ at $x = a$ has then the following interpretations:

1. In general, it measures the instantaneous rate of change of the value of the variable $f(x)$ with respect to the variable x, when the latter has the value a.

2. In the special geometric interpretation of the equation $y = f(x)$ as a curve, it measures the slope of the curve at the point $[a, f(a)]$; the equation of the line tangent to the curve at this point is $y - f(a) = f'(a)(x - a)$.

These remarks in combination with what has preceded carry some important consequences:

I. If for a value a of x, the derivative $D_x f(x)$ is positive, i.e. if $f'(a) > 0$, the function $f(x)$ is on the increase; if the derivative is negative, the function is on the decrease; and conversely.

II. The values of x for which the function $f(x)$ attains a maximum or minimum must be looked for among the numbers which satisfy the condition $D_x f(x) = 0$.

III. If with every value of x, we associate the value at that point of the derivative with respect to x of a given function $f(x)$, we obtain a new function called the *first derived function* of $f(x)$; it is usually denoted by $f'(x)$.

IV. The derived function $f'(x)$ subjected to the treatment that was applied on the preceding pages to the function $f(x)$, gives rise

to the *second derived function*, usually denoted by $f''(x)$. We can proceed in this manner, and obtain derived functions of successively higher orders.

It will have become evident to the reader that the concepts to which he has been introduced sweep over a very extensive domain. In each of the scientific fields in which functional relations have been established (compare p. 305), the derived functions have a special interpretation of significance in that field. It is no exaggeration therefore if we say that the general concept under which all these interpretations can be subsumed is the source from which comes insight into a wide range of facts and phenomena. In the remainder of this chapter this should become increasingly apparent, even though, for obvious reasons, we shall have to confine ourselves to a restricted portion of this range. We shall be able to follow only one or two of the many streamlets which contribute to the river whose course we are surveying.

136. Up hills and down dales. The calculation of the derivatives of functions, called differentiation, can in general be made by direct application of Definition XLIX. The labor necessary for this work is very much reduced by means of some general principles, based on the theory of limits. The development of these principles belongs to the *Differential Calculus*. A large number of textbooks, long, short and medium-sized; good, bad and indifferent; expensive, cheap and moderately priced, have been published on this subject. No interested reader should have the slightest difficulty in finding one to suit his needs and tastes.

Our programme does not include the technique of differentiation. In accordance with our general plan, we are concerned with the general concepts and with the manner in which they operate; incidentally we hope to reach the objective set up as the aim of this expedition. Such tools as may be needed on the way will be prepared as we go along (compare also *138*).

One of the most important among the simple applications of the derivative was alluded to under II above. We will therefore go a little further into the question of the maximum and minimum values of a function.

Let us consider first a simple example, viz. the problem so to divide 25 into two parts, that their product shall be a maximum. If one part be denoted by x, then the other part is $25 - x$. The question under consideration is therefore that of determining a

value of the variable x such that the corresponding value of the function $x(25 - x)$, i.e. the function $25x - x^2$, shall be a maximum. In accordance with our earlier result we can say that such a value must be looked for among those values of x for which the derivative of the function $f(x) = 25x - x^2$ is equal to 0. Now it is easily found [1] that $D_x(25x - x^2) = 25 - 2x$, so that there is only one value of x, viz. $12\frac{1}{2}$ for which this derivative takes the value 0. For this value of x the function takes the value $156\frac{1}{4}$; is this the maximum value of the function?

Before this question can be answered we must make clear what we mean by a *maximum value* and a *minimum value* of a function. Ordinarily these words are used to designate a largest and smallest, a most and a least; and it is not always specified very carefully among what collection of competitors it is most or least. For our present purpose, we have to be satisfied with the following definition:

Definition L. The value $f(a)$ is called a *maximum* (or *minimum*) value of the function $f(x)$ if we can indicate an interval of values of x containing a in its interior such that if a_1 is any other value in that interval, then $f(a_1) < f(a)$ (or $f(a_1) > f(a)$).

Nothing is said about the size of this interval; it may be very small. If we think of the graphical interpretation of the equation $y = f(x)$, it seems clear intuitively that if the derivative exists at a maximum or minimum point, it must be equal to 0 there. On the other hand, a function need not have either a maximum or a minimum at every point a at which $f'(a) = 0$; this fact is illustrated by the origin in Fig. 74. A point on a curve such as the origin on the locus of $y = x^3$ in that diagram is called a *point of horizontal inflection*. A careful study of the question shows that maximum points, minimum points and points of horizontal inflection are the only points at which the first derivative can be equal to 0; these possibilities are illustrated in Fig. 76.

This figure also brings out the fact that a maximum (minimum) value of a function $f(x)$ in the sense of Definition L need not be the largest (least) value of the function. To distinguish between largest (least) value, and maximum (minimum) values of a function in the sense of that definition, the term absolute maximum (minimum) is used to designate the former and the term relative maximum (minimum) to indicate the latter. A function may have several

[1] Compare *138*, 5.

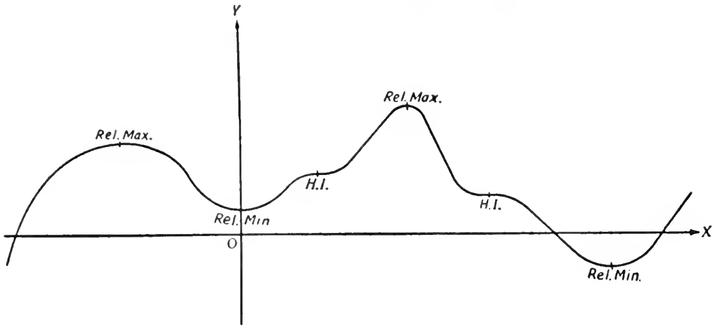


FIG. 76

relative maxima (minima), or none, but it can have only one absolute maximum (minimum). It can also happen that a relative maximum (minimum) is at the same time the absolute maximum (minimum), as in the example treated in this section. We shall be concerned chiefly with relative maxima and minima and shall therefore omit the word "relative" when this can be done without risk of misunderstanding.

Returning to our special problem, we recall that $D_x f(x) = 0$ only if $x = 12\frac{1}{2}$, and that $f(12\frac{1}{2}) = 156\frac{1}{4}$. If we observe now that

$$\begin{aligned} f(12\tfrac{1}{2} + h) &= 25 \cdot (12\tfrac{1}{2} + h) - (12\tfrac{1}{2} + h)^2 \\ &= 25 \cdot 12\tfrac{1}{2} + 25h - (12\tfrac{1}{2})^2 - 25h - h^2 \\ &= 156\tfrac{1}{4} - h^2, \end{aligned}$$

it follows that $f(12\frac{1}{2} + h) < f(12\frac{1}{2})$, for any positive or negative value of h. Hence we conclude that $156\frac{1}{4}$ is, both relatively and absolutely, the maximum value of the product of the two parts into which 25 can be divided.

137. The stream supplies power. What can we say about the general problem of maxima and minima of a function $f(x)$? In view of what has gone before, it is clear that we have to find a way of distinguishing between the three kinds of points illustrated in Fig. 76. The points of maximum ordinate, of minimum ordinate and of horizontal inflection have this property in common, that at each of them the slope of the curve passes through the value zero. Now it is well known that a varying magnitude may pass through zero in four different ways: (1) the magnitude may come from positive values and go to negative values; (2) it may be on its

way from negative to positive values; (3) it may reach zero from positive (negative) values and immediately return to positive (negative) values. In the first case the magnitude decreases through zero, in the second it increases through zero, in the third and fourth cases it skims zero. If we consider now the three kinds of points illustrated in Fig. 76, and if we have clearly in mind the contents of Theorem LXIV (p. 294) and of *134*, 11, we see that as the independent variable x increases, (1) at a maximum point the slope of the curve *decreases* through zero, (2) at a minimum point the slope of the curve *increases* through zero, (3) at a point of horizontal inflection the slope *skims* zero. Next we apply I on page 317; we can then distinguish between the three types of points as follows:

maximum point: derivative of slope is negative;
minimum point: derivative of slope is positive;
point of horizontal inflection: derivative of slope is zero.[1]

But since the slope of a curve is determined by the derivative of the corresponding function, the derivative of the slope of a curve is given by the derivative of the derivative of the corresponding function, i.e. by the values of its second derived function. We obtain therefore the following results (remember IV on p. 317):

maximum point: $f'(x) = 0$ and $f''(x) < 0$;
minimum point: $f'(x) = 0$ and $f''(x) > 0$;
point of hor. inflection: $f'(x) = 0$ and $f''(x) = 0$.

The difficulty hinted at in the last footnote arises from the fact that when the second derived function $f''(x)$ passes through zero, each of the three possibilities mentioned on page 320 can arise; thus further and more detailed considerations are necessary. The present analysis although obviously incomplete should suffice to indicate the method of procedure.

In the problem discussed in *136*, we found that $f(x) = 25x - x^2$, that $f'(x) = 25 - 2x$ and that $12\frac{1}{2}$ is the only value of x for which $f'(x) = 0$. It follows very easily that $f''(x) = -2$, so that the slope is steadily decreasing as the independent variable x increases;

[1] As a matter of fact the conclusions stated here and in the next paragraph are not correct. The situation is a little more involved than becomes apparent. Readers may try to discover for themselves where the argument is weak; they may consult a good book on the Calculus such as, e.g., Hardy, *Pure Mathematics* (1925), pp. 220, 268.

this is in accord with our earlier conclusion that $156\frac{1}{4}$ is indeed a maximum value of the function.

Let us now consider a few other examples:

1. From a rectangular piece of cardboard, square corners are cut out (see Fig. 77); then the sides are turned up about the lines AB, BC, CD and DA so as to form a box. If the sides of the piece of cardboard are 12 and 15 inches, how large a square should be cut out to obtain the box of maximum content?

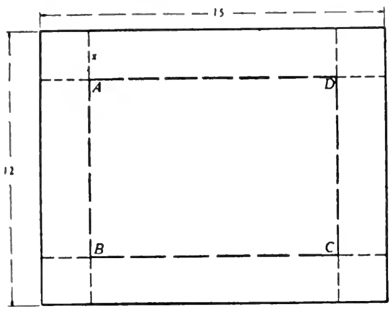


FIG. 77

Let us denote the side of the square to be cut out, the independent variable, by x; and the volume of the box by y. The dimensions of the box are then $15 - 2x$, $12 - 2x$ and x; consequently the relation between x and y is given by $y = x(15 - 2x)(12 - 2x)$. It is our problem to determine the value of x for which y has a maximum. A simple calculation gives $y = 4x^3 - 54x^2 + 180x$; of this function we must calculate the 1st and 2nd derivatives. We proceed as on pages 312–316.

Since $f(x) = 4x^3 - 54x^2 + 180x$, we find that

$$f(x + h) = 4(x + h)^3 - 54(x + h)^2 + 180(x + h).$$

Hence

$$f(x + h) - f(x) = 12x^2h + 12xh^2 + 4h^3 - 108xh - 54h^2 + 180h;$$

and

$$\frac{f(x + h) - f(x)}{h} = 12x^2 + 12xh + 4h^2 - 108x - 54h + 180.$$

Therefore

$$\begin{aligned} f'(x) = D_xf(x) &= \lim_{h \to 0} \frac{f(x + h) - f(x)}{h} \\ &= 12x^2 - 108x + 180.\text{[1]} \end{aligned}$$

[1] See the footnote on p. 314.

Application of the same process to this new function gives:

$$f'(x+h) - f'(x) = 12(x+h)^2 - 108(x+h) + 180 - (12x^2 - 108x + 180)$$
$$= 24xh + 12h^2 - 108h,$$

so that

$$f''(x) = \lim_{h \to 0} \frac{f'(x+h) - f'(x)}{h}$$
$$= \lim_{h \to 0} (24x + 12h - 108) = 24x - 108.$$

The values of x for which the first derived function vanishes are found by solving the quadratic equation

$$12x^2 - 108x + 180 = 0,$$

which reduces to $x^2 - 9x + 15 = 0$. Its solutions are $\frac{9+\sqrt{21}}{2}$ and $\frac{9-\sqrt{21}}{2}$.[1] If we calculate $f''(x)$ for each of these values we find that

$$f''\left(\frac{9+\sqrt{21}}{2}\right) = 12\sqrt{21} > 0, \text{ and } f''\left(\frac{9-\sqrt{21}}{2}\right) = -12\sqrt{21} < 0.$$

Hence the side of the square to which corresponds the box of largest volume is $\frac{9-\sqrt{21}}{2} = 4.5 - 2.2913 \cdots = 2.2087 \ldots$. The volume of this box is obtained by substituting $\frac{9-\sqrt{21}}{2}$ for x in the function $4x^3 - 54x^2 + 180x$; we find that the maximum volume is equal to $81 + 21\sqrt{21} = 177.2341 \ldots$. Is this the absolute maximum?

If we tabulate the values of $f(x)$ and of $f'(x)$ for some values of x near the one which furnishes the maximum, we obtain the following results:

x	0	1	2	2.1	2.2	2.3	2.4	2.5	3
$f(x)$	0	130	176	176.90	177.23	177.01	176.26	175	162
$f'(x)$	180	84	12	6.12	.48	− 4.92	− 10.08	− 15	− 36

[1] Compare p. 133, formulas (7.4).

The labor involved in calculating this table is justified at this point in our study only because it brings out clearly the fact that when $f'(x)$ is large and positive (as at the start), $f(x)$ increases rapidly with x. As x tends towards the value $\dfrac{9-\sqrt{21}}{2} = 2.2087\ldots$, $f'(x)$ decreases towards zero and $f(x)$, although still increasing, grows much more slowly. After the critical value has been passed $f'(x)$ becomes negative, and $f(x)$ decreases, the more rapidly the larger the numerical value of $f'(x)$ becomes. These facts are brought out very clearly on a graphical representation of the equation $y = 4x^3 - 54x^2 + 180x$.

2. A wire, 25 inches in length, is to be divided into two parts. One of these is to be bent into a circular shape, the other into a square. How large should the parts be in order that the combined area of the square and the circle shall be as large as possible?

The independent variable x is the length of one of the pieces; let us take it to be that of the piece to be bent into circular shape. The radius of that circle will then be equal to $\dfrac{x}{2\pi}$ and its area will be $\pi\left(\dfrac{x}{2\pi}\right)^2 = \dfrac{x^2}{4\pi}$ sq. in. The other piece of wire, of length $25 - x$,

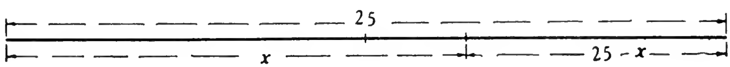


FIG. 78

is then bent into a square whose sides are each equal to $\dfrac{25-x}{4}$ and whose area is therefore equal to $\dfrac{(25-x)^2}{16}$. The problem before us then is to determine a value of x such that the function $y = \dfrac{x^2}{4\pi} + \dfrac{(25-x)^2}{16}$ shall have a maximum value. The method of attack should now be familiar. From

$$\begin{aligned} f(x) &= \frac{x^2}{4\pi} + \frac{(25-x)^2}{16} \\ &= \left(\frac{1}{4\pi} + \frac{1}{16}\right)x^2 - \frac{25x}{8} + \frac{25^2}{16}, \end{aligned}$$

we find

$$f(x + h) = \left(\frac{1}{4\pi} + \frac{1}{16}\right)(x^2 + 2hx + h^2) - \frac{25(x + h)}{8} + \frac{25^2}{16},$$

so that

$$f(x + h) - f(x) = \left(\frac{1}{4\pi} + \frac{1}{16}\right)(2hx + h^2) - \frac{25h}{8},$$

and

$$\frac{f(x + h) - f(x)}{h} = \left(\frac{1}{2\pi} + \frac{1}{8}\right)x - \frac{25}{8} + \left(\frac{1}{4\pi} + \frac{1}{16}\right)h.$$

Therefore

$$f'(x) = \operatorname*{Lim}_{h \to 0} \frac{f(x + h) - f(x)}{h} = \left(\frac{1}{2\pi} + \frac{1}{8}\right)x - \frac{25}{8}.$$

Similarly, we find

$$f''(x) = \frac{1}{2\pi} + \frac{1}{8}.$$

To solve our problem we have to determine a value of x for which $f'(x) = 0$; there is only one such value, viz.

$$x = \frac{\frac{25}{8}}{\frac{1}{2\pi} + \frac{1}{8}} = \frac{25\pi}{4 + \pi}.$$

But, for this value of x, $f''(x) > 0$; indeed the second derived function is positive for all values of x. We conclude that the function $f(x)$ occurring in this problem has no relative maximum (compare p. 319), but that for $x = \frac{25\pi}{4 + \pi} = 11$ in. (approximately), the total area of circle and square has a relative minimum.

Our question has not yet been answered. For there remains a possibility that the function $f(x) = \left(\frac{1}{4\pi} + \frac{1}{16}\right)x^2 - \frac{25x}{8} + \frac{25^2}{16}$ has an absolute maximum which is not a relative maximum, when x varies over the range of values from 0 to 25, the only ones which have significance in the problem. To investigate this possibility, we apply I, page 317, to the information which we have secured concerning $f(x)$. From the fact that $f'(x) = \left(\frac{1}{2\pi} + \frac{1}{8}\right)x - \frac{25}{8}$,

we have already concluded that $f'(x) = 0$ if and only if $x = \dfrac{25\pi}{4+\pi}$, let us say if $x = 11$. But we know furthermore that $f'(x)$ will be positive if $x > 11$ and negative when $x < 11$. This means that $f(x)$ increases steadily once x has passed beyond 11, and also when x

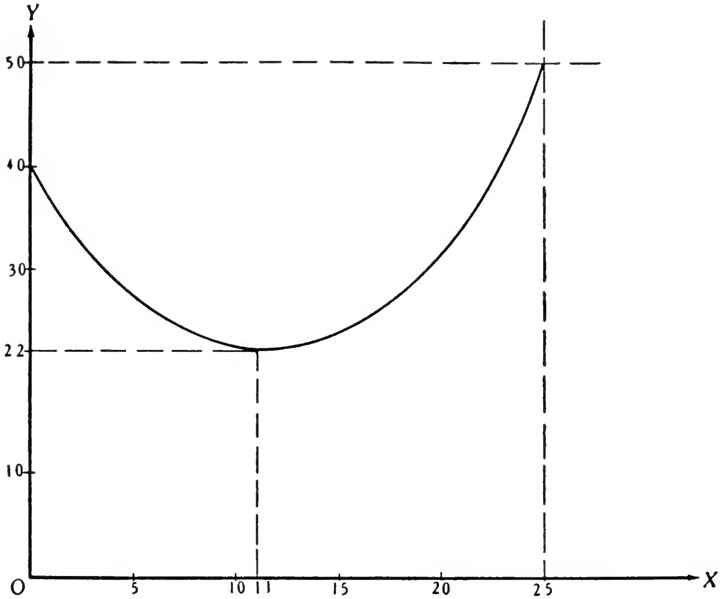


FIG. 79

decreases away from 11. The general character of the locus of the equation $y = f(x)$ is therefore indicated by Fig. 79.

It follows that our function $f(x)$ does have an absolute maximum — it must be either $f(0)$ or $f(25)$. A simple calculation shows that $f(0) = \dfrac{25^2}{16}$ and that $f(25) = \dfrac{25^2}{4\pi}$; but $4\pi < 16$ and therefore $f(25)$ is the maximum value. We conclude that the maximum value of the combined areas of square and circle is attained when the entire wire is bent into circular shape; if *two* pieces must be produced, then there is no maximum for their combined areas. Incidentally

we have found that the minimum obtained by making $x = \dfrac{25\pi}{4 + \pi}$ is an absolute as well as a relative minimum.

138. Operating the machinery.

1. Determine the first derived function of each of the following functions: (a) $2x^2$; (b) $7x^2$; (c) ax^2, where a is independent of the value of x.

2. Determine the equation of the tangent line to the curve $y = x^3$ at the points $A(-2, -8)$ and $B(2, 8)$. What can be said about these two lines?

3. Use the method developed in the preceding sections to find the derivative of the function $\dfrac{1}{x}$ at the point $(1, 1)$.

4. Determine the first and second derived functions of the function $\dfrac{1}{x}$. What property of this function is brought out by the first derived function?

5. Calculate the first and the second derived functions of the function $ax^2 + bx + c$ in which a, b and c are independent of the value of x.

6. Use the results of 5 to prove that the function $ax^2 + bx + c$ always possesses either a relative maximum or a relative minimum, but never both, provided $a \neq 0$.

7. Use the results of 5 and 6 to represent graphically the general character of the function $ax^2 + bx + c$.

8. Determine the first derived function of the function x^4. (The details in procedure do not differ essentially from those in earlier questions of this sort.)

9. Determine the first derived function of the function $\sqrt{x}$. (Remember that $\sqrt{x}$ represents the non-negative number whose square is x and that only non-negative values of x are to be considered. A new technical device is needed at one point in the application of the general method to this function.)

10. Obtain the first derived function of $\sqrt{x}$ by using the fact that if $y = \sqrt{x}$, then $x = y^2$.

11. Use the method of 10 to determine the first derived function of $\sqrt[3]{x}$; verify the conclusion by the direct method.

12. Find the equation of the tangent line to the curve $y = \dfrac{1}{x}$ at an arbitrary point $\left(a, \dfrac{1}{a}\right)$ on the curve; indicate the general character of the curve.

13. Show that the right triangle formed by the coördinate axes and the tangent line to the curve $y = \frac{1}{x}$ at an arbitrary point P has an area which is independent of the choice of the point P.

14. An open rectangular box is to be constructed with square base. Its volume is to be 500 cubic inches. What should be its dimensions in order that the amount of material needed for its construction be a minimum?

15. Solve the same problem as the one given in 14, but for a closed box.

16. We have again to construct a rectangular box with square base and top. Its content is to be 2500 cubic inches. The material used for the bottom is to cost 3 times as much as that used for the sides; the material for the top costs twice as much as that for the sides. What are the most economical dimensions?

17. Set up and solve problems similar to those given in 14, 15 and 16, but involving a cylindrical box.

18. Determine the derivative of a variable with respect to itself. What is the geometric significance of this problem?

19. Determine the derivative of a constant c (i.e. a magnitude whose value is the same for all values of the variable x) with respect to x.

20. Prove that the sum of two infinitesimals is an infinitesimal.

21. Prove that the statements: "x tends to the limit a" and "$x - a$ is an infinitesimal" are equivalent statements.

22. Prove that if for each of the functions $u(x)$ and $v(x)$ the first derived functions exist, then the same is true for the function $w(x) = u(x) + v(x)$, and $w'(x) = u'(x) + v'(x)$.

23. Prove that the product of an infinitesimal by a constant is also an infinitesimal.

24. Prove that if the first derived function of $u(x)$ exists and c is a constant, then the first derived function of $w(x) = c \cdot u(x)$ also exists and $w'(x) = c \cdot u'(x)$.

25. Prove that the product of two infinitesimals is an infinitesimal.

26. Prove that if the first derived functions of $u(x)$ and of $v(x)$ exist, then the first derived function of $w(x) = u(x) \cdot v(x)$ exists as well, and $w'(x) = u(x) \cdot v'(x) + u'(x) \cdot v(x)$.

27. Prove that under the hypotheses of 26, the first derived function of $w(x) = \frac{u(x)}{v(x)}$ exists and is equal to $\frac{v(x) \cdot u'(x) - u(x) \cdot v'(x)}{v^2(x)}$, provided $v(x) \neq 0$.

139. An airplane view. The preceding sections should have given the reader an idea of some of the problems to which the differential calculus gives access. They have dealt but very little with the technique of the subject. This involves in the first place

the determination of the derived functions of all the functions which arise in elementary work, such as the rational functions, the trigonometric functions, the logarithms and exponential functions (compare *144*, *156*, *157*). It is in connection with the derivative of the logarithmic function, that the base e of the Napierian logarithms makes its most natural appearance (see *144*, IV). A knowledge of the derivatives of those various functions widens of course the scope of possible applications to maxima and minima problems.

Along with such applications comes a fuller understanding and broader use of the concept "rate of change of one variable with respect to another variable." It has already been observed (compare p. 316) that the derivative of $f(x)$ with respect to x measures the instantaneous rate of change of the variable $f(x)$ with respect to the variable x. This interpretation of the derivative gets especial significance when x is the measure of time; in that case it becomes "*the* rate of change." For when we speak of *the* rate of change of a magnitude we always mean its rate of change with respect to *time* as independent variable. If y measures the distance of a point moving along a straight line l from a fixed point O on that line and x is a measure of time, the change in y which corresponds to a change h in x means the distance through which the point moves in time h. Hence the ratio of the change in y to the change in x designates the *average velocity* or average speed of the moving point during the time h. The limit of this ratio, as $h \to 0$, i.e. the derivative of y with respect to x, measures therefore the (*instantaneous*) *velocity* of the moving point. Moreover, if y' measures the velocity of such a moving point, the ratio of the change in y' to the change h in x measures the average change in velocity, i.e. the *average acceleration* during the time h. The second derivative of y with respect to x is then the (*instantaneous*) *acceleration* of the moving point. If we use the little knowledge of differential calculus which we have acquired we can conclude that if a point moves along a line in such a way that its distance, y, from a fixed point is expressed in terms of the time x elapsed since a fixed moment of time by the formula $y = 2 + 3x - 5x^2$, its velocity at time x will be equal to $3 - 10x$ and its acceleration will be constantly equal to -10.

In a similar way, the speed with which water flows out of a reservoir means the rate of change with respect to time of the

amount of water that has come out of the reservoir (or, in numerical value, of the amount of water that has remained in); the speed with which a train moves along a straight track is the rate of change with respect to time of the distance of the train from some fixed point along the track. If we know that the amount of water in the reservoir after time t is equal to w, and that $w = 500 - 2t - t^2$ gallons, then the speed of flow is equal to $2 + 2t$; etc.

There are many other important instances of the derivative of one variable with respect to another. We have already recognized the slope of a curve as the derivative of the ordinate with respect to the abscissa, i.e. as the limit towards which tends the ratio of the *increase in the ordinate* to the *increase in the abscissa* as the latter tends to zero, provided of course this ratio does approach a limit (compare *129* and *135*). As a point describes a simple curve, many variables occur besides the ordinate and the abscissa of the moving point, e.g. the direction of the curve (i.e. the angle which the tangent line to the curve makes with the X-axis), the slope of the curve (i.e. the tangent of that angle), the length of the curve measured from some fixed point on it, the area enclosed by the curve, a fixed ordinate, the X-axis and the varying ordinate. It is of value and interest to consider the derivative of any one of these variables with respect to any other one. For example, the derivative of the slope with respect to the abscissa, i.e. the second derivative of y with respect to x has proved to be of importance in distinguishing between maximum points, minimum points and points of horizontal inflection.

Another of these possible derivatives is the "curvature" of a curve, the concept which started us on our present expedition, and which we are now prepared to discuss with greater advantage. The curvature of a simple plane curve is the rate of change of the direction of the curve with respect to the length of its arc. To reduce this definition to simple terms, we proceed in exactly the same way as in *135* where the derivative of the ordinate with respect to the abscissa was introduced.

We suppose that we have a simple plane curve at every point of which a tangent line can be drawn. We draw the tangent line $T'T$ to the curve at the fixed point A (see Fig. 80) and also at the successive points B_1, B_2, . . . which tend towards A along the curve. The points of intersection of the tangent line at A with the tangent lines at B_1, B_2, . . . are denoted by S_1, S_2, ...

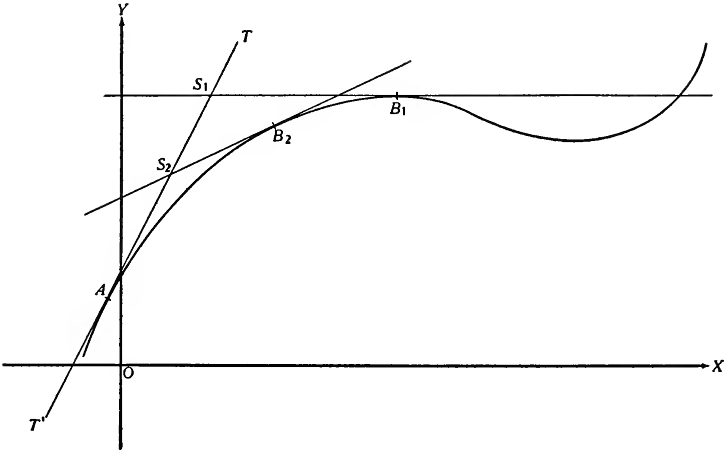


FIG. 80

Then the change in the direction of the curve as we move along the curve from B_1 to A is measured by the angle B_1S_1T; the change in direction from B_2 to A is measured by the angle B_2S_2T, . . .; the change in direction from B_n to A is measured by the angle B_nS_nT. If we divide the measures of these angles by the lengths of the arcs of the curve from A to B_1, to B_2, . . . to B_n we obtain the "average curvature" of these arcs. The curvature of the curve at A is, in accordance with our definitions, the limit towards which the ratio $\dfrac{\text{measure of } \angle B_nS_nT}{\text{length of arc } AB_n}$ tends as B_n approaches A along the curve.[1]

Here we have an answer to the question as to what is to be understood by the curvature of a curve (see *126*). There still remains the problem of determining a formula which will enable us to calculate the curvature of a simple curve when its equation is given in the form $y = f(x)$. This problem is one of those dealt with in the differential calculus. Since the details of the derivation of this formula require a good deal of preliminary work, we shall content ourselves with stating that the result is $\dfrac{y''}{(1 + y'^2)^{\frac{3}{2}}}$. For

[1] It is to be understood that the angle B_nS_nT is the numerically least angle through which S_nB_n can be rotated to bring it into coincidence with S_nT.

example, for the curve $y = x^2$, for which $y' = 2x$ and $y'' = 2$ the curvature is equal to $\frac{2}{(1 + 4x^2)^{\frac{3}{2}}}$; this shows that the curve has the greatest curvature at the point where $x = 0$. The denominator of the fraction $\frac{y''}{(1 + y'^2)^{\frac{3}{2}}}$ which gives the curvature is always positive (compare Def. XXIV, p. 104); the numerator may be positive, zero or negative. Therefore the curvature of a curve will turn out to be positive, zero or negative, according as y'' has one of these signs. The geometrical significance of these three possibilities will become clear if we repeat for an arbitrary point of the curve the analysis made on pages 320–322 for points at

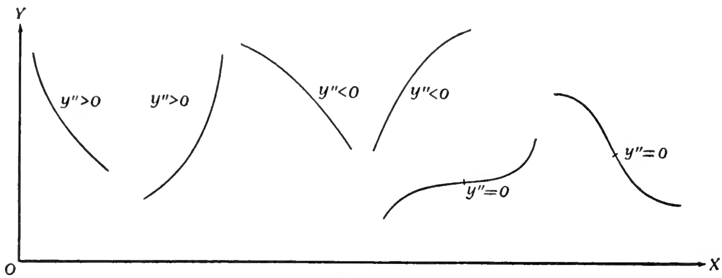


FIG. 81

which $y' = 0$. We shall leave this analysis to the reader and shall merely state the result, viz. at a point at which $y'' > 0$, the curve has the shape of a hollow bowl turned upwards (*concave upward* is the technical term); when $y'' < 0$, the shape is that of a bowl turned downwards (*concave downward*); when $y'' = 0$ the curve changes from one of these shapes to the other (such a point is called a point of inflection); compare Fig. 81. We see that the three types of points discussed on pages 320–321 are special cases of the possibilities indicated here.

To summarize the discussion of the curvature of a curve and to give the reader a sample of the technique of the differential calculus, we shall calculate the curvature of a circle.

We know that $x^2 + y^2 = r^2$ is the functional relation between the abscissa x and the ordinate y of points on the circle of radius r

whose center is at the origin (compare p. 303). From this relation we have to calculate y' and y'', and then the function $\dfrac{y''}{(1 + y'^2)^{\frac{3}{2}}}$.

(a) To obtain y' we proceed as follows: If $A(x, y)$ and $B(x + h, y + k)$ are two points on the circle, the slope of the line AB is equal to $\dfrac{k}{h}$ (Theorem LXVIII); and y', which measures the slope of the line tangent to the circle at A, is equal to the limit towards which $\dfrac{k}{h}$ tends as $h \to 0$. Since B and A lie on the circle, we have

$$(12.7) \qquad (x + h)^2 + (y + k)^2 = r^2 \text{ and } x^2 + y^2 = r^2.$$

Subtracting the second of these equations from the first we are led successively to the equations

$$\begin{aligned} 2hx + h^2 + 2ky + k^2 &= 0, \\ k(2y + k) &= -h(2x + h), \end{aligned}$$

and

$$\frac{k}{h} = -\frac{2x + h}{2y + k}.$$

We draw now on the assumption made in *135* (see footnote on p. 316), which guarantees that $k \to 0$ as $h \to 0$. We find then that the slope of the circle at the point A is given by

$$(12.8) \qquad y' = \operatorname*{Lim}_{h \to 0} \frac{k}{h} = -\frac{x}{y}.$$

We stop for a moment to observe (1) that this formula shows that the slope of the circle fails to exist at the points where $y = 0$; (2) that it shows that the circle has positive slope in the 2nd and 4th quadrants, and negative slope in the 1st and 3rd quadrants (compare p. 317, I); (3) that, if combined with *130*, 6 and 11, it shows that the tangent line to the circle at A is perpendicular to the radius OA.

(b) To calculate y'' we could simply use *138*, 27 and 18; but this would be too easy. Let us therefore suppose that $y' + i$ is the slope of the circle at B; then $y'' = \operatorname*{Lim}_{h \to 0} \dfrac{i}{h}$. Since (12.8) is applicable for any point on the circle, we have

$$y' + i = -\frac{x+h}{y+k} \quad \text{and} \quad y' = -\frac{x}{y}.$$

Therefore

$$i = -\frac{x+h}{y+k} + \frac{x}{y} = \frac{kx - hy}{y(y+k)},$$

and

$$\frac{i}{h} = \frac{\dfrac{xk}{h} - y}{y(y+k)}.$$

The theorems mentioned in the footnote on page 314, and one additional theorem, now come into play. By means of them, and by use of (12.8) and (12.7), we find then that

$$(12.9) \qquad y'' = \frac{xy' - y}{y^2} = -\frac{x^2 + y^2}{y^3} = -\frac{r^2}{y^3}.$$

This result shows that the upper semi-circle on which $y > 0$, is concave downward, while the lower semi-circle is concave upward.

(c) From (12.8), we derive by using also (12.7) that

$$(1 + y'^2)^{\frac{3}{2}} = \left(\frac{x^2 + y^2}{y^2}\right)^{\frac{3}{2}} = \frac{r^3}{|\,y^3\,|}.$$ [1]

This last formula, in conjunction with (12.9), leads to the conclusion that the curvature of the upper semi-circle is equal to $-\dfrac{1}{r}$ and that of the lower semi-circle to $\dfrac{1}{r}$. Thus we have found some interesting results:

1. The numerical value of the curvature is the same at all points of a given circle; thus we can speak of the "curvature of a circle."

2. The curvature of a circle is inversely proportional to its radius; i.e. the smaller the radius, the greater the curvature, and inversely.

3. The curvature of the upper half of a circle is negative, that of the lower half is positive.

These conclusions bring out the fact that the definition of the

[1] Remember that $(1 + y'^2)^{\frac{3}{2}}$ is always positive, and recall the notation $|\,a\,|$ mentioned on p. 295.

curvature of a curve which was given on page 330, is in accord, at least for the circle, with the intuitive understanding of the term. The mathematical treatment has to that extent proved to be a usable abstraction from more concrete experience.

At last we can explain in a more satisfactory way than before the meaning of the terms positive and negative curvature used in the preceding chapter to describe the character of a surface in the neighborhood of one of its points.

Arbitrary planes through a point P on a surface S will cut out curves on the surface which pass through P. To every curve that can be obtained in this way we draw the tangent line at P.[1] All these tangent lines lie in one plane, called the tangent plane to the surface S at P. The line through P perpendicular to this tangent plane is called the normal to the surface at P.

We considered in *124* the curves obtained as intersections with planes through the normal to S at P; these are called the normal sections of the surface at P. For each of the curves obtained in this way, the curvature at P can be determined by the method described in the preceding paragraphs. If these curvatures all turn out to be of one sign (all positive or all negative), we say that the surface S has positive curvature at P; if it happens that some of these curves have positive curvature, while others have negative curvature at P, we say that the surface S has negative curvature at that point. An examination of Fig. 60 shows that this represents a surface for which at every point all the normal sections are either concave downward, or else all concave upward; in either case the curvatures of the normal sections at any one point are all of one sign; the surface has therefore positive curvature at every point. A surface of this kind is said to be of constantly positive curvature. The saddle surface shown in Fig. 61 has at every point some normal sections which are concave upward and some which are concave downward; it has negative curvature at every point and is called a surface of constantly negative curvature.

140. The removal of an obstacle. A further remark is necessary before we can bring this chapter to a close. In the various examples of derivatives which have been discussed, we had to determine the

[1] This statement clearly involves the assumption that all these curves have a slope at P (compare Definition XLV) and thus indirectly an assumption as to the character of the surface. The analysis of this assumption lies beyond our present range; it is dealt with in books on the Theory of Surfaces. We have to refer to these books also for a proof of the statement in the next sentence above.

limit of a ratio of the form $\frac{k}{h}$ as $h \to 0$. For some of them it was obvious that $k \to 0$ as $h \to 0$; for the others this fact was explicitly assumed (compare the footnotes in *135*; further examples can be furnished by the reader who has tried to operate the machinery in *138*). We want now to free ourselves of this assumption; we can do so, at least in part, by showing that if the derived function exists at all, then the variable k is indeed an infinitesimal. For, let L be the limit of $\frac{k}{h}$ as $h \to 0$. Then we know that the numerical value of $\frac{k}{h} - L$ is an infinitesimal (compare Definitions XLIII and XLIIIa; also *138*, 21); let us denote it by α. It follows then that $\frac{k}{h} - L = \pm \alpha$ and hence that $k = \pm h\alpha + hL$. From this equation, we conclude by means of *138*, 23, 25 and 20, that k is an infinitesimal.

A derivative is therefore the limit of the ratio of two infinitesimals, which may be related to each other in various ways. The central problem of the differential calculus consists then in the determination of such limits; it is a part of the wider domain known as the infinitesimal calculus.[1] In particular cases the determination of the limit of the ratio of two infinitesimals may turn out to be far from easy; the few examples which we have taken up have been selected in such a way as to avoid complications. They have perhaps been sufficient to make clear that also in the realm of infinitesimals there are differences of "size," paradoxical as this may seem to "common sense." The situation has some similarity with the one we discovered in the study of infinite sets (compare Chapter II) among which we found an unexpected variety of possibilities. If the reader has had his interest in this subject aroused, he can find ample opportunity to satisfy it in numerous books on analysis. In the next chapter we shall turn to another part of the infinitesimal calculus.

[1] It may not be superfluous to remind the reader of the fact that this does *not* mean the calculus of the very small, but the calculus of variables which approach 0.

CHAPTER XIII

EX PARVIS COMPONERE MAGNA

Had I been told when I was at the university that others understood the integral calculus while I did not, that would have touched my pride. — Levin, in Tolstoy's *Anna Karenina*.

141. A moving wall. We consider now in greater detail one of the derivatives connected with a plane curve of which the possibility was suggested on page 330, viz. that of the area enclosed by the curve, a fixed ordinate AA', the X-axis and the varying ordi-

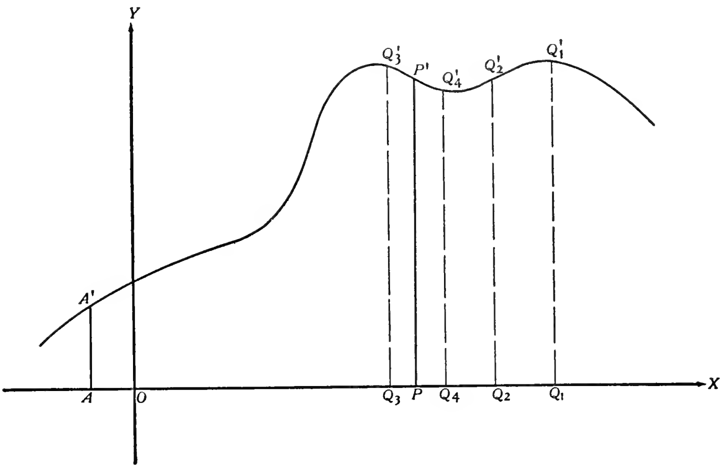


FIG. 82

nate PP', with respect to the abscissa (compare Fig. 82). It turns out that this derivative can be calculated very easily; moreover the method which we shall use for its calculation is of interest in itself.

In accordance with Definition XLIX, the derivative in which we are interested is the limit approached by the ratio of the change

in the area $A'APP'$ which results from a change in the distance AP (or, what amounts to the same, a change in the abscissa OP) to this change in AP, as the latter change tends to zero. To calculate this limit, let us take a succession of points on the X-axis, Q_1, Q_2, . . . Q_n, . . . through which we draw ordinates $Q_1Q'_1$, $Q_2Q'_2$, . . . $Q_nQ'_n$, ... It is to be understood that, as n increases, the point Q_n tends towards P, and the corresponding ordinate $Q_nQ'_n$ will tend towards the ordinate PP'.[1] The changes in the area $A'APP'$ which result from these changes in the abscissa OP are the areas $P'PQ_1Q'_1$, $P'PQ_2Q'_2$, . . . $P'PQ_nQ'_n$, . . . all of which would be trapezoids except for the fact that $P'Q'_1$, $P'Q'_2$, . . . $P'Q'_n$, . . . are not straight lines, but arcs of the curve. We have therefore to consider the sequence of ratios $\frac{\text{area } P'PQ_1Q'_1}{PQ_1}$, $\frac{\text{area } P'PQ_2Q'_2}{PQ_2}$, . . . $\frac{\text{area } P'PQ_nQ'_n}{PQ_n}$, ...[2]

How can we decide whether this sequence tends towards a limit, and how can we determine the limit in case one does exist? The situation is rather different from those we have met in earlier examples; it calls for a new idea. This we proceed to develop.

142. The principle of the fly-swatter. Suppose we have three variables, each running through its own set of real values; let us call them x, y, z. Let us suppose moreover that these variables are related to each other in such a way, that with each value of any one of them is associated a definite value of each of the others. We shall then be concerned with the case in which, if $\bar{x}$, $\bar{y}$, $\bar{z}$ is a set of associated values of these variables, $\bar{y}$ will always be intermediate in value to $\bar{x}$ and $\bar{z}$.[3] We shall express this fact by saying that the variable y is between the variables x and z; in formula we shall write $x \leqq y \leqq z$ or $x \geqq y \geqq z$.

[1] While we have specified what is to be understood by the phrase "the point Q_n tends towards P" (compare p. 295), it is not immediately evident what is meant by "the *curve* C_n tends towards the curve C." The ideas involved in this phrase are of considerable importance, but beyond our scope and purpose. For understanding what is to be the interpretation of "the corresponding ordinate $Q_nQ'_n$ will tend towards the ordinate PP'" (which is the only instance of the approach of curves that concerns us now), we shall rely upon the reader's intuitive grasp of the meaning of these words; this is most likely to guide him aright.

[2] In Fig. 82, PQ_1, PQ_2, PQ_4 and the corresponding areas are positive, but PQ_3 and the area $P'PQ_3Q'_3$ are negative; all the ratios indicated above are positive. For a curve which lies partly below the X-axis, the analogous ratios may vary in sign.

[3] Here it is to be understood that if $\bar{x} = \bar{z}$, then we shall also have $\bar{y} = \bar{z}$.

Examples of such variables can readily be found in experience. If I try to catch a moth between my two hands, the distances, from some fixed plane of the moth and of the feet of the perpendiculars from the moth to my hands, satisfy the condition on the variables y, x and z respectively which we have in mind. The reader will no doubt recognize other examples in his own experience.

Concerning such variables we shall now prove an important theorem, viz.

Theorem LXXI. If x, y and z are three real variables, so related that $x \leqq y \leqq z$ or $x \geqq y \geqq z$, and if both x and z tend to the limit L, then y also tends to the limit L.

Proof.[1] We have to have in mind clearly the content of Definition XLIII (see p. 295) and the discussion following it. In that case we will know that the variation of x and z is of such nature that $|L - x|$ and $|L - z|$ become and remain less than e.[2] But this carries with it that $|x - z|$ will become less than $2e$; consequently since y is always between x and z, the numerical value of $x - y$ becomes and remains less than $2e$. Now $L - y = (L - x) + (x - y)$. From this we conclude that the numerical value of $L - y$ does not exceed the sum of the numerical values of $L - x$ and of $x - y$. The former of these becomes and remains less than e, the latter less than $2e$; hence $|L - y|$ becomes and remains less than $3e$. If we clearly understand the meaning of the symbol e, we conclude from this that y tends to the limit L, as was to be proved (compare *138*, 20).

Interesting special cases of this theorem arise when y, or x, or z is a constant. The first of these special cases will come up in a later connection. The others lead to the following statement.

Corollary. If two variables, s and t, and a constant L are so related that $L \leqq s \leqq t$, or so that $L \geqq s \geqq t$ and if t approaches L as a limit, then s also tends to L.

We shall frequently refer to this simple and useful theorem by a name suggested by the corollary, viz. the fly-swatter principle.

[1] In the study of this proof and of later proofs concerning real variables, it will be helpful to draw a diagram, making use of the assumption concerning the isomorphism between the system of real numbers and the system of points on a directed straight line, which was made and discussed in Chapter V (compare pp. 72–75). In this isomorphism the numerical value of the difference between two real numbers corresponds to the undirected distance between the two points associated with these numbers.

[2] Compare pp. 295, 296; geometrically this means that the "point x" and the "point z" lie in an interval of width $2e$ of which the center is at the "point L."

Its first service consists in getting us over the difficulty with which we were confronted at the end of *141* (see also *145*, 7 and *147*).

Let us consider, in connection with the area $P'PQ_nQ'_n$ of Fig. 82 (reproduced separately in Fig. 83), the largest rectangle of base PQ_n which is entirely contained in it; we shall denote its area by m_n. Let us also consider the smallest rectangle of base PQ_n *in* which this area $P'PQ_nQ'_n$ is entirely contained; and let us denote its area by M_n. As Q_n approaches P the numbers m_n, area $P'PQ_nQ'_n$

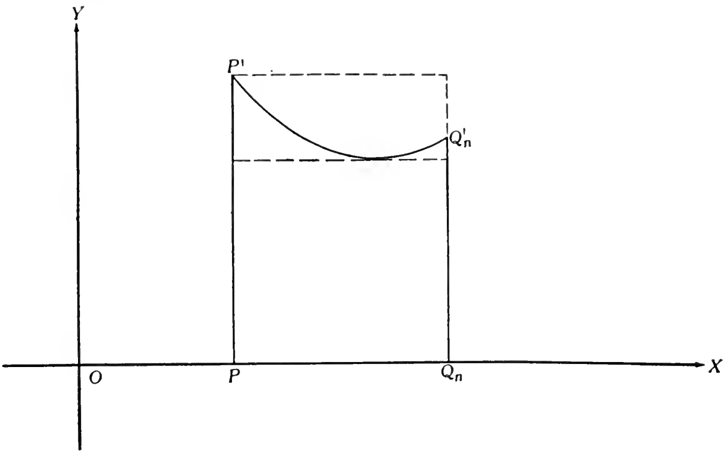


FIG. 83

and M_n vary; and they vary in such manner that $m_n \leqq$ area $P'PQ_nQ'_n \leqq M_n$. Consequently the variables $\frac{m_n}{PQ_n}$, $\frac{\text{area } P'PQ_nQ'_n}{PQ_n}$ and $\frac{M_n}{PQ_n}$ are conditioned by the inequalities

$$\frac{m_n}{PQ_n} \leqq \frac{\text{area } P'PQ_nQ'_n}{PQ_n} \leqq \frac{M_n}{PQ_n},$$

if PQ_n is positive and by the reversed inequalities if PQ_n is negative. But this shows that the variable $\frac{\text{area } P'PQ_nQ'_n}{PQ_n}$ is between the variables m_n and M_n; hence the first condition of the "fly-swatter principle" is satisfied. But the second condition is also satisfied.

For $\frac{m_n}{PQ_n}$ = the least ordinate between PP' and QQ'_n; and $\frac{M_n}{PQ_n}$ = the largest ordinate between PP' and QQ'_n. Therefore as Q_n tends towards P, both $\frac{m_n}{PQ_n}$ and $\frac{M_n}{PQ_n}$ approach the length of the ordinate PP' as a limit. Theorem LXXI is therefore applicable; hence we conclude that, as Q_n tends towards P, the ratio $\frac{\text{area } P'PQ_nQ'_n}{PQ_n}$ tends towards the limit PP'. This is the answer to the question asked in the last paragraph of *141*. We formulate our conclusions as follows:

Theorem LXXII. If $A(x)$ denotes the area enclosed by the curve whose equation is $y = f(x)$, a fixed ordinate, the X-axis and the variable ordinate which corresponds to the variable abscissa x, then $D_xA(x)$ is equal to $f(x)$; i.e. $A(x)$ is a function whose first derived function is $f(x)$.

Example. Take the curve whose equation is $y = x^2$. If the shaded area OPP' contained between this curve, the positive X-axis and a variable ordinate PP' be denoted by $A(x)$ (compare Fig. 84), we conclude that $D_xA(x) = x^2$.

Can we use this result to determine the area $A(x)$? Here we have a problem after the manner of Jacobi. For it is clearly the inverse of the problem with which we were occupied in the preceding chapter. Instead of being given a function and being required to find its derivative, we are now given the derived function and we are asked to determine the function of which it is the derivative. The technique in differential calculus which we have acquired is not sufficient to let us go very far in the solution of this inverse problem in its general form. But it may be just about adequate to enable us to deal with this particular example. For, we should know from *135* (b) and *138*, 2 that $3x^2$ is the derived function of x^3; and hence that $\frac{x^3}{3}$ has for its derivative the given function x^2. Moreover, this function $\frac{x^3}{3}$ vanishes when $x = 0$, as it should if it is to represent the variable area OPP'. We conclude therefore that area $OPP' = \frac{x^3}{3}$. It is interesting to observe that, since the ordinate

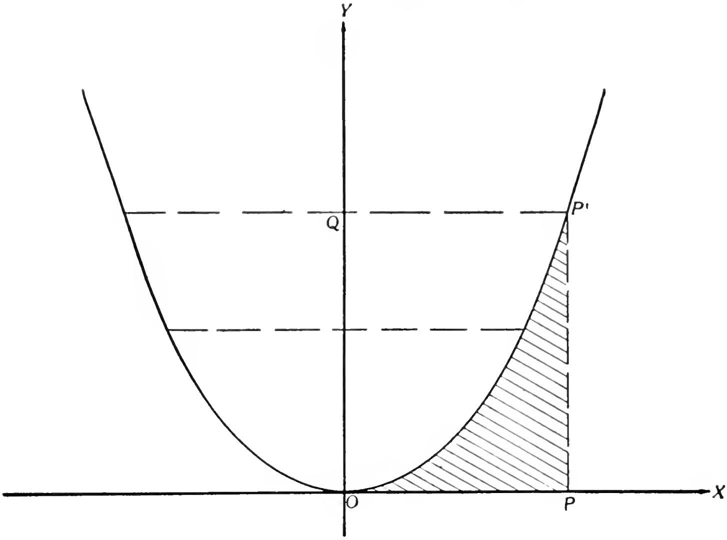


FIG. 84

PP' is equal to x^2 and therefore the area of the rectangle $QOPP'$ equal to x^3, the fact that area $OPP' = \dfrac{x^3}{3}$ shows that the curve $y = x^2$ divides this rectangle into two parts, equal in area to one third and two thirds of that rectangle respectively.

143. The determination of areas. The application which we have made of Theorem LXXII in this particular example suggests a general method for determining areas bounded by curves. This alluring vista we must pursue a little further; but, some additional equipment is required for this venture.

Theorem LXXIII. If the derived function of a function $u(x)$ is equal to 0 for all values of x in an interval (ab) then $u(x)$ is constant throughout that interval.

A proof of this theorem is a bigger job than we can undertake here. What we can do, however, is to make the conclusion appear plausible; and this is the purpose of the next paragraph. The reader who remains dissatisfied with such treatment (all praise to him!) can readily find solace in good books on the infinitesimal calculus, or in books on the Theory of Functions of a Real Variable.

Assume that there were two points on the interval (ab) for which the corresponding values of $u(x)$ were different from each other; in particular, suppose that $u(x_1) < u(x_2)$ (see Fig. 85). Then somewhere on the way from x_1 to x_2, the function $u(x)$ must be increasing. Hence, since the hypothesis of the theorem tacitly supposes that the derivative of $u(x)$ exists at all points on the interval (ab),

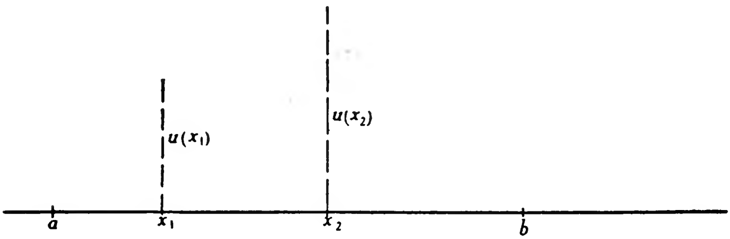


Fig. 85

there must be at least one point on (ab) at which $u'(x)$ is positive (compare *135*, I); but this contradicts our hypothesis. We conclude that no two values of x between a and b can give rise to different values of $u(x)$, i.e. that $u(x)$ = constant.

Remark. The converse of this theorem can readily be proved (see *138*, 19).

Theorem LXXIV. If $U'(x) = u(x)$, then any function of the form $U(x)$ + constant has $u(x)$ as its first derived function; conversely, any function whose first derived function is $u(x)$ will have the form $U(x)$ + constant.

Proof. The first half of this theorem is an immediate consequence of *138*, 22 and 19. The second half follows from the former of these, in conjunction with *138*, 24 and Theorem LXXIII. For, if $U'(x) = u(x)$ and $U'_1(x) = u(x)$, then $U'_1(x) - U'(x) = 0$, and hence $[U_1(x) - U(x)]' = 0$; consequently $U_1(x) - U(x)$ = constant, or $U_1(x) = U(x)$ + constant.

Suppose now that we have a curve determined by an equation $y = f(x)$ and that we wish to determine the area bounded by it, the X-axis, and the ordinates which correspond to the abscissas a and b (see Fig. 86). We introduce a variable ordinate, corresponding to the general abscissa x and consider the area bounded by the curve, the X-axis, the line $x = a$ and this variable ordinate; we shall denote it again by $A(x)$. We know from Theorem LXXII

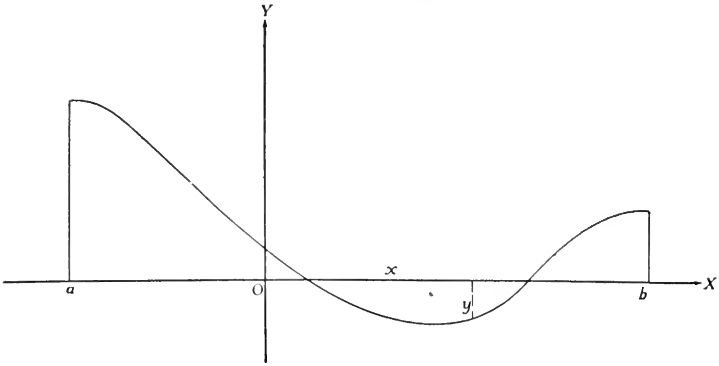


FIG. 86

that $A'(x) = f(x)$. *Suppose now that we can find a function $F(x)$ whose first derived function is $f(x)$.* Then we conclude on the basis of Theorem LXXIV that $A(x)$ has the form $F(x)$ + constant, i.e.

$$(13.1) \qquad A(x) = F(x) + \text{constant}.$$

This constant is as yet undetermined. But we can easily determine it if we remember that the left boundary of the area represented by $A(x)$ is the line $x = a$. For this carries with it that the area vanishes when its right boundary is also formed by this ordinate corresponding to the abscissa a, i.e. that

$$A(a) = 0, \text{ or that } F(a) + \text{constant} = 0.$$

If we combine this relation with (13.1) we conclude that

$$A(x) = F(x) - F(a).$$

Finally, the required area, bounded on the right by the ordinate for $x = b$, is given by $A(b)$, i.e. by $F(b) - F(a)$. We have the following important result:

Theorem LXXV. The area bounded by the curve $y = f(x)$, the X-axis and the ordinates which correspond to the abscissa a and b is equal to $F(b) - F(a)$, where $F(x)$ is *any function whose derivative is equal to $f(x)$.*

Remark. Theorem LXXII was obtained under the supposition that the abscissas increase towards the right. It follows that areas above the X-axis are counted as positive if described from left to

right; and that areas below the X-axis are counted as negative if described in that sense (compare footnote 2 on page 338).

The theorem just proved opens up the possibility of finding areas bounded by arbitrary curves. The actual realization of this possibility depends upon the ability to find a function whose derivative is known. Something is gained if we have a simple way of describing the thing we are looking for — hence a definition.

Definition LI. Any function whose derivative with respect to x is equal to a given function $u(x)$ is called *an indefinite integral* of $u(x)$ with respect to x; it is denoted by the symbol $\int u(x)$.

In this terminology, we can phrase the content of Theorem LXXV as follows:

Theorem LXXVa. The area bounded by the curve $y = f(x)$, the X-axis and the ordinates $x = a$ and $x = b$, is equal to the difference between the values for $x = b$ and for $x = a$ of any indefinite integral of $f(x)$.

For the phrase "the difference between the values for $x = b$ and for $x = a$ of any indefinite integral of $f(x)$," we use the form: "the indefinite integral of $f(x)$ between the limits a and b," and the symbolism $\int_a^b f(x)$. Thus, if $F(x)$ is any indefinite integral of $f(x)$, we have $\int_a^b f(x) = F(b) - F(a)$; the area described in theorems LXXV and LXXVa is therefore represented by $\int_a^b f(x)$.

The determination of indefinite integrals of various functions constitutes the first part of the *integral calculus*. It clearly requires a good technique in differential calculus to become skillful in this part of the integral calculus. We can not hope to get very far in walking on our hands until we have acquired enough stability to stand on our feet. In general we can not expect much success in inverting until we have something to invert. But we must try to get all we can out of the little technique in differential calculus which we did acquire; incidentally we may strengthen this technique a little. In the meantime we shall have to accept, largely without proof, a few general statements which will give some freedom of action in the integral calculus.

144. A foraging expedition. The first point of importance is the following:

I. The derivative with respect to x of the indefinite integral with respect to x of any function $u(x)$ is equal to $u(x)$, i.e.

$$D_x \int u(x) = u(x); \tag{13.2}$$

moreover,

$$\int D_x u(x) = u(x) + \text{constant}, \tag{13.2a}$$

i.e. a function whose derivative with respect to x is equal to $D_x u(x)$ can differ from $u(x)$ at most by a constant.

These formulas express the contents of Definition LI and of Theorem LXXIV; they can be used at least to verify some of the results which follow.

II. The indefinite integral of the sum or difference of two functions is equal to the sum or difference of their indefinite integrals.[1] (Compare *138*, 22.)

III. The indefinite integral of x^0 or of any positive integral power of x is the next higher integral power, divided by the exponent of this higher power, i.e. if $n \geqq 0$, $\int x^n = \dfrac{x^{n+1}}{n+1} + \text{constant}$. In particular, $\int x^2 = \dfrac{x^3}{3} + \text{constant}$, $\int x^3 = \dfrac{x^4}{4} + \text{constant}$, etc.

To prove this statement, it will obviously be sufficient, in view of (13.2) and (13.2*a*), to show that $D_x x^n = nx^{n-1}$ for any natural number n. Any one who knows the binomial theorem can easily prove this formula by the direct method used in Chapter XII. The following proof is based on preceding results and on *138*, 26. For we have already seen that the formula is true for $n = 1, 2, 3, 4$ (compare *138*, 8 and 18). Suppose now that it holds for $n = 1, 2, \ldots k$. Then

$$\begin{aligned} D_x x^{k+1} = D_x(x^k \cdot x) = x^k D_x x + x D_x x^k &= x^k + x \cdot k \cdot x^{k-1} \\ &= x^k + kx^k = (k+1)x^k; \end{aligned}$$

this shows that the formula still holds for $n = k + 1$. The principle of mathematical induction now completes the proof.

IV. The indefinite integral of $\dfrac{1}{x}$ is equal to $\log_e x + \text{const.}$, where

[1] We omit the words "with respect to x" when there is no danger of confusion.

e is the Napierian base (compare p. 329). We already know that this statement is equivalent to the formula $D_x \log_e x = \frac{1}{x}$. It is regrettable that we cannot go in for a proof of this very interesting and important formula. Its importance will become evident from the frequent references to it which occur in the sequel; its interest arises in part from the fact that it introduces the base of the natural logarithms which we have met on previous occasions. But once more our program compels us to abandon, however reluctantly, a very tempting excursion. The interested reader, etc.

V. The indefinite integral of any negative integral power of x, except x^{-1}, is equal to the power of x whose exponent is one more than the exponent of the given power, divided by the new exponent; i.e. if $n > 0$ and $\neq 1$, we have $\int x^{-n} = \frac{x^{-n+1}}{-n+1} + \text{const.}$ This formula can be deduced from I and III by means of *138*, *27*.

VI. The indefinite integral of a power of x, whose exponent is any rational number p, different from -1 is equal to the power of x with an exponent $p+1$, divided by the new exponent; i.e. if p is a rational number, $\neq -1$, then $\int x^p = \frac{x^{p+1}}{p+1} + \text{const.}$ It is clear that VI contains III and V as special cases. To prove its additional content, we would have to show that if $\frac{n}{m}$ is any rational number in which $m \neq 1$, then $D_x x^{\frac{n}{m}} = \frac{n x^{\frac{n}{m}-1}}{m}$. This can be done without difficulty by use of the binominal theorem if we follow the method used on pages 333 and 334; the details will be omitted.

Combinations of these theorems will enable the reader to carry out a good many "elementary integrations."

It follows from Definition LI in conjunction with the discussion on page 329 that the indefinite integral can also be used to determine the velocity of a point moving in a line if its acceleration is given as a function of the time; and the distance traversed by such a point if its velocity is given as a function of the time. For the complete determination of the velocity and the position, certain initial conditions would also have to be known. There are many other problems for which the indefinite integral provides a solution; they are discussed in books on the integral calculus.

145. Working with borrowed tools.

1. Find the area bordered by the line $y = 2x - 1$, the X-axis and the lines $x = -1$ and $x = +2$.

2. Determine indefinite integrals of the following functions:

$$\text{(a) } 3x^2 - 4x + 2; \quad \text{(b) } x^2 - 2 + \frac{1}{x^2}; \quad \text{(c) } \sqrt{x} - 2\sqrt[3]{x}.$$

3. Determine the area between the curve $y = \sqrt{x}$ (see *138*, 9), the X-axis and the line $x = 4$.

4. Determine the areas of each of the three parts into which the square formed by the points (0,0), (0,1), (1,1) and (1,0) is divided by the two curves $y = x^2$ and $y = \sqrt{x}$.

5. Solve the problem obtained from the one stated in 4, by using the curves $y = x^3$ and $y = \sqrt[3]{x}$ instead of the curves there mentioned.

6. Determine the area bounded by the Y-axis, the curve $y = \sqrt[3]{x}$ and the line $y = 5$.

7. Prove that if two real variables x and y are so related to each other and to the constant C, that $C \leqq x \leqq y$, and if y approaches C as a limit, then x also approaches the limit C.

8. Use the method followed in the proof of *144*, III, to show $D_x(ax + b)^n = an(ax + b)^{n-1}$, when a and b are independent of x, and n is a natural number.

9. Determine indefinite integrals with respect to x of each of the following functions: (a) $(x - 2)^3$; (b) $(2x + 3)^2$; (c) $\left(\frac{x}{3} + 1\right)^4$.

10. Assuming that, similarly to the extension from *144*, III to *144*, VI, the formula proved in 8 remains valid if the exponent is any rational number p different from -1, obtain indefinite integrals with respect to x of the functions

$$\text{(a) } \sqrt{x + 2}; \quad \text{(b) } \sqrt[3]{2x - 4}; \quad \text{(c) } \sqrt{\frac{2x}{3} - 1}; \quad \text{(d) } (3x + 2)^{-2}.$$

11. Show that the area bounded by the curve $y = x^2$, the X-axis and the line $x = 4$ is equal to the area enclosed between the curve $y = (x - 3)^2$, the X-axis and the line $x = 7$.

12. Prove that the indefinite integral of the function $\sqrt{x}$ between the limits 0 and 4 is equal to the indefinite integral of the function $\sqrt{x + 3}$ between the limits -3 and 1.

13. Prove that in general $\displaystyle\int_a^b f(x) \equiv \int_{a+k}^{b+k} f(x - k)$.

14. Prove that $D_x(x^2 + a)^n = 2nx(x^2 + a)^{n-1}$, when a is independent of x and n is a natural number. (The method of *144*, III and of 8 can be used to advantage!)

15. Develop a formula for $D_x(x^3 + a)^n$, with the agreements of 14.

16. Assuming that the formula proved in 14 is still valid for an arbitrary rational exponent $p, \neq -1$, determine indefinite integrals with respect to x of the following functions: (a) $x(x^2 + 2)^{-2}$; (b) $x\sqrt{x^2 - 3}$; (c) $x(x^2 - 3)^3$.

17. Prove that $\int uD_xv + \int vD_xu = uv$, when u and v are functions of x.

18. The acceleration of a body moving in a straight line is 4 ft. per sec. per sec. At time $t = 0$, the velocity is 0, i.e. the body is at rest. Determine the velocity after 5 seconds.

19. A body falling in a vacuum is subject to a constant downward acceleration of approximately 32 ft. per sec. per sec. (this is known as the acceleration of gravity). If a body falls from rest at time $t = 0$, what will be its velocity after 4 seconds?

20. Solve the problem of 19 if at time $t = 0$ the body had an upward velocity of 70 ft. per sec.

21. Returning to the body of 18, how far from the position it occupied at time $t = 0$ would this body be after 10 seconds?

22. If a body were dropped (not thrown) from a point in a vacuum 400 ft. above the ground at time $t = 0$, where would it be and what would be its velocity after 3 sec.? How long would it take to reach the ground?

23. At time $t = 0$ a body is thrown down with a velocity of 15 ft. per sec. What would be its velocity after 5 sec.? How far below its starting point would it be at that time?

24. A ball is thrown vertically upward with a speed of 48 ft. per sec. What would be its speed after 1 sec.; after 2 seconds?

25. How long would this ball continue to rise and what height would it reach?

26. How long after being thrown up would the ball be back at its starting point?

27. Denoting the constant acceleration due to gravity by g, determine the velocity of a falling body at time t after it starts from rest.

28. Determine the velocity of the falling body at time t if its velocity at time $t = 0$ is equal to v_0.

29. Find the distance from the starting point at time t of the falling body whose velocity at the start is equal to v_0.

146. Making an estimate. The great significance of the integral calculus rests upon the fact that it deals with properties of configurations taken as a whole. The derivative $u'(x)$ of a function $u(x)$ gives us information as to a certain characteristic of the

corresponding curve *at and near one of its points;* it tells us the speed of a moving body *at one moment.* The first and second derivatives together give us information concerning the curvature of a curve *at one point*, or about the acceleration *at one moment* of a moving body. On the other hand the indefinite integral of a function $u(x)$ enables us to determine the area of an entire region, or the distance traveled by a moving body throughout an interval of time. This contrast is frequently expressed by the statement, that the differential calculus deals with *properties in the small*, whereas the integral calculus deals with *properties in the large.* The integral calculus enables one to find volumes and areas of solids, weights of bodies of uniform or of non-uniform densities, forces of attraction exerted by bodies of various shapes, and so forth. We must get some idea as to how this is done and how the concepts with which we have become acquainted are used in the solution of such problems.

For this purpose we shall first treat in some detail a very special problem, viz. that of determining the gravitational attraction, exerted by a vertical bar upon a point outside. Suppose, in order not to complicate matters unduly, that the line AB represents a bar of length $2l$; that is to say, let us neglect the thickness and breadth of the bar (see Fig. 87). Let us furthermore suppose that the bar is of uniform density, which means that any two equally long segments of the bar contain equal amounts of matter. We wish to determine the gravitational attraction exercised by this bar upon a unit of mass placed at a point P, situated at a distance d from the bar AB, on its perpendicular bisector.

The gravitational attraction *between two masses*, m_1 and m_2, concentrated in two points a distance a apart is equal to $\dfrac{Rm_1m_2}{a^2}$, wherein R is a constant about which we shall not be concerned, and which we shall take equal to 1 in the sequel. This formula is not directly applicable to our problem because we are dealing with two masses of which only one is concentrated at a point, which is at unequal distances from different parts of the other mass. It is in a situation of this kind that the integral calculus intervenes. The bar AB is thought of as divided into $2n$ equal parts; let l_k, $k = 1, \ldots 2n$ be the length of an arbitrary one of the parts. Let the vector f_k denote the attraction which this part exerts on the point P. Then the total attraction is equal to the resultant of the forces

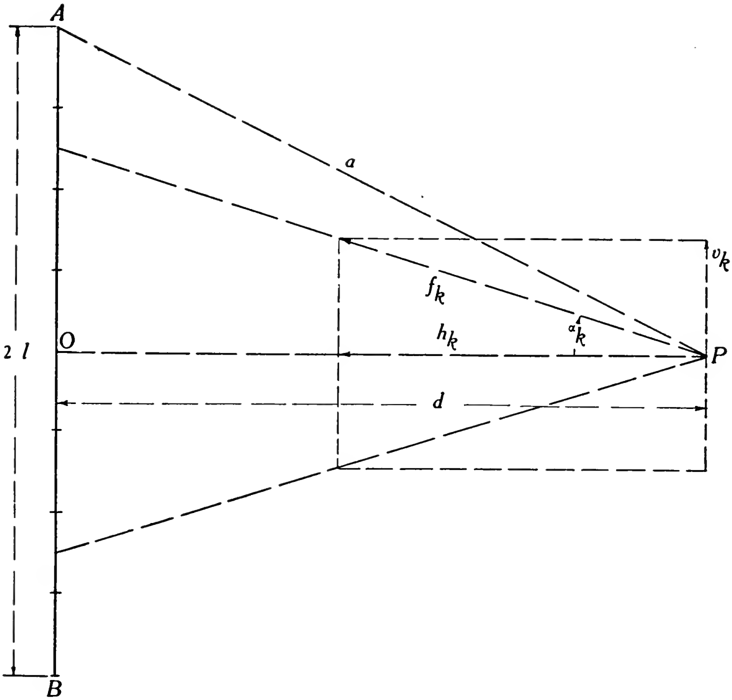


FIG. 87

f_k, $k = 1, \ldots n$. To determine this resultant, we resolve each force f_k in a vertical component v_k and a horizontal component h_k.

Of the vertical components we can dispose readily. For, to each part of the bar above the middle of the bar, can be made to correspond an equal part equally far below the middle. The vertical components of the attractions exercised by these two parts are equal in size, but opposite in direction; hence they neutralize each other. Moreover, the horizontal components of the attractions due to these two parts of the bar are equal in magnitude *and* in direction. It follows that the resultant of all the vertical components vanishes and that we need be concerned only with the horizontal components of the upper half of the bar.

If we denote by α_k the angle which the force f_k makes with the line PO, we see that the horizontal component of f_k is equal to

$|f_k| \cos \alpha_k$ so that the total attraction which we are seeking to determine is a force in the direction PO whose magnitude is given by the sum $2[|f_1| \cos \alpha_1 + |f_2| \cos \alpha_2 + \cdots |f_n| \cos \alpha_n]$. It remains to determine this sum; but we can not do this directly because neither the magnitudes of the forces f_k, exerted by the different parts, nor the angles α_k which determine their directions are any more readily found than would be the force exerted by the entire bar. Suppose now that the entire mass of the part l_k were concentrated at the upper boundary P_k of that part (see Fig. 88); the attractive force which this mass would then exert would be less

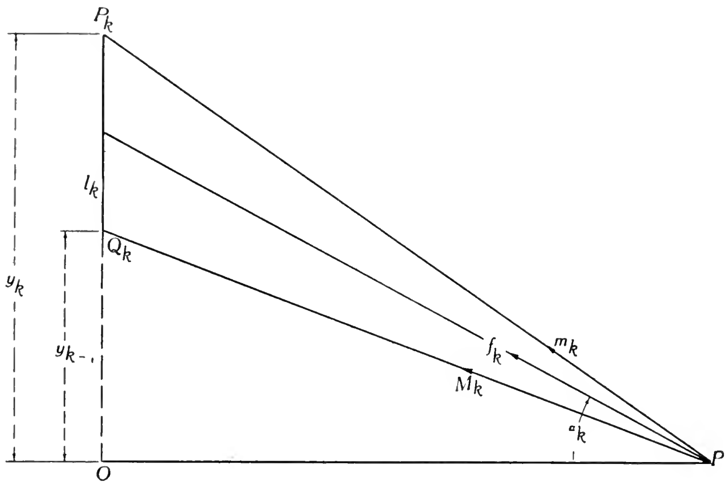


FIG. 88

than it actually is; let us call it m_k. Moreover the horizontal component of this fictitious force would be a smaller fractional part of $|m_k|$, than the horizontal component h_k is of $|f_k|$, for the angle OPP_k is greater than α_k, so that $\cos OPP_k < \cos \alpha_k$ (compare pp. 98–101).

If, on the other hand, the mass of the part l_k were concentrated at the lower end Q_k, the horizontal component of the force would be greater than it actually is; and that on two counts, for the force would be increased in magnitude (we shall call it M_k) and its angle with OP would be diminished, so that the cosine of this angle would be increased.

Let us denote now by y_k the distance from O to P_k. Then, since the lower end, Q_k, of the part l_k coincides with the upper end P_{k-1} of the part l_{k-1}, the mass of the part l_k is equal to $(y_k - y_{k-1})\delta$, if δ is the uniform density of the bar. The magnitude of the force m_k is then equal to $(y_k - y_{k-1})\delta \div \overline{PP_k}^2$; and $\cos OPP_k = \dfrac{d}{PP_k}$, while $\overline{PP_k}^2 = d^2 + y_k^2$.

We conclude that the horizontal component of the fictitious force m_k which would occur if the mass were concentrated at P_k is equal to $\dfrac{(y_k - y_{k-1})\delta d}{(d^2 + y_k^2)^{\frac{3}{2}}}$. By similar reasoning we find that the horizontal component of M_k would be equal to $\dfrac{(y_k - y_{k-1})\delta d}{(d^2 + y_{k-1}^2)^{\frac{3}{2}}}$. Hence we have the following fundamental inequality:

$$(13.3) \qquad \frac{(y_k - y_{k-1})\delta d}{(d^2 + y_k^2)^{\frac{3}{2}}} \leqq |f_k| \cos \alpha_k \leqq \frac{(y_k - y_{k-1})\delta d}{(d^2 + y_{k-1}^2)^{\frac{3}{2}}}.$$

From this it follows that the total force we are trying to determine will lie between the sum of the terms like the one on the left in (13.3) and the sum of those like the one on the right. The successive terms in these sums are obtained if we replace the subscript k which occurs in them in turn by the integers $1, 2, 3, \ldots n-1, n$. To indicate such a sum of terms which differ only in the number to be substituted for a variable subscript, we use the notation $\sum\limits_{k=1}^{n}$. Thus we have, if F designates the force of attraction of the bar on the point P:

$$(13.4) \qquad 2\sum_{k=1}^{n} \frac{\delta d(y_k - y_{k-1})}{(d^2 + y_k^2)^{\frac{3}{2}}} \leqq |F| \leqq 2\sum_{k=1}^{n} \frac{\delta d(y_k - y_{k-1})}{(d^2 + y_{k-1}^2)^{\frac{3}{2}}}.$$

This result constitutes the first step in the determination of the required force by the method of the integral calculus.

147. Tightening the screws. The second step is a little difficult to carry out; it is usually omitted from the elementary books on the calculus. It consists of two parts; first we have to determine whether the sums on the right and on the left in (13.4) tend towards a limit as n, the number of equal parts into which the upper part OA of the bar is divided, is allowed to increase indefinitely. We

shall have to leave this question open, at least for the present; before we finish the treatment of our problem we shall come a little nearer to answering it.

In preparation for the second part, we shall prove a theorem on limits closely related to the Corollary of Theorem LXXI.

Theorem LXXVI. When two real variables, x and y, vary in such manner that (1) their difference is an infinitesimal and (2) one of them approaches a limit L, then the other one also tends to L.

Proof.[1] Let us suppose that x tends towards the limit L. Then the two parts of the hypothesis tell us (1) that $|x - y| < e$ and[2] (2) that $|\mathrm{L} - x| < e$. Now $\mathrm{L} - y = (\mathrm{L} - x) + (x - y)$; hence $|\mathrm{L} - y| \leqq |\mathrm{L} - x| + |x - y| < 2e$. This shows that $\mathrm{L} - y$ is an infinitesimal; in other words, that y tends towards L, as was to be proved.

As suggested on page 339, this theorem is but a special case of Theorem LXXI; it should also be clear that the Corollary of the theorem (the fly-swatter in the exact sense) can be deduced from this special case.

This theorem fits exactly into the problem we have before us. In place of the two variables x and y we have the sums which appear on the left and right of (13.4). Let us calculate their difference. Barring the factor δd which is the same for every term in both sums, we find that these sums are respectively equal to

$$x = \frac{y_1 - y_0}{(d^2 + y_1{}^2)^{\frac{3}{2}}} + \frac{y_2 - y_1}{(d^2 + y_2{}^2)^{\frac{3}{2}}} + \cdots + \frac{y_{n-1} - y_{n-2}}{(d^2 + y_{n-1}{}^2)^{\frac{3}{2}}} + \frac{y_n - y_{n-1}}{(d^2 + y_n{}^2)^{\frac{3}{2}}};$$

and

$$y = \frac{y_1 - y_0}{(d^2 + y_0{}^2)^{\frac{3}{2}}} + \frac{y_2 - y_1}{(d^2 + y_1{}^2{}_2)^{\frac{3}{2}}} + \cdots + \frac{y_{n-1} - y_{n-2}}{(d^2 + y_{n-2}{}^2)^{\frac{3}{2}}} + \frac{y_n - y_{n-1}}{(d^2 + y_{n-1}{}^2)^{\frac{3}{2}}}.$$

If we remember now that the parts into which AB was divided were *equal parts*, i.e. that $y_1 - y_0 = y_2 - y_1 = \cdots = y_{n-1} - y_{n-2} = y_n - y_{n-1} = \frac{l}{n}$, we can conclude that the second term in the second sum

[1] See footnote 1 on p. 339.

[2] Statements like $|x - y| < e$ in this proof are to be understood to mean that x and y vary in such manner that $|x - y|$ becomes and remains less than the arbitrary positive number e; compare the proof of Theorem LXXI, p. 339.

equals the first term in the first, the third term in the second equals the second in the first, etc., the last term in the second equals the next to the last term in the first; in general, the kth term in the second sum equals the $(k - 1)$th term in the first sum, for $k = 2, \ldots n$. Therefore if we subtract the first sum from the second, there will remain only the first term in the second sum minus the last term in the first; i.e., if we observe moreover that $y_0 = 0$ and $y_n = l$:

$$y - x = \frac{y_1 - y_0}{d^3} - \frac{y_n - y_{n-1}}{(d^2 + l^2)^{\frac{3}{2}}}.$$

Let us now replace $y_1 - y_0$ and $y_n - y_{n-1}$ by their common value, viz. $\frac{l}{n}$; we find then that

$$y - x = \frac{l[(d^2 + l^2)^{\frac{3}{2}} - d^3]}{d^3(d^2 + l^2)^{\frac{3}{2}}n}.$$

In this expression for $y - x$ everything is fixed in value except the factor n in the denominator. Hence we can conclude that as n increases indefinitely $|y - x|$ will tend to 0, i.e. that $y - x$ is an infinitesimal.

What can we deduce from this conclusion? By means of Theorem LXXVI we can evidently conclude that, if either of the sums in (13.4) tends towards a limit, then the other sum will tend towards the same limit. But if this is the case, we can apply the "fly-swatter principle," Theorem LXXI; and this would tell us that $|F|$ which lies between the two sums must tend towards this same limit as well. On the other hand, if we allow our physical intuition to play a part in this game, we know that the attractive force F, which we are trying to determine, has a fixed magnitude, independent of the number of parts into which we conceive the bar to be divided; i.e. F is a constant, independent of n. Thus the assumption that at least one of the variable sums in (13.4) tends to a limit leads to a conclusion which is in accord with physical intuition. This agreement is frequently claimed as a support for the validity of the assumption. It is a nice question of epistemology to find out on what grounds the claim is justifiable; it is patently extraneous to a purely mathematical treatment. This requires the further study to which we alluded at the opening of the present section, and which we must forego. Subject to this limitation we have

then proved that the magnitude of the attraction which we are looking for is given by

$$(13.5) \qquad |F| = 2 \lim_{n \to \infty} \sum_{k=1}^{n} \frac{\delta d(y_k - y_{k-1})}{(d^2 + y_k^2)^{\frac{3}{2}}}.\text{[1]}$$

This completes the second step in the solution of our problem.

148. Achievement. Before proceeding to the third step, we had best introduce another definition. A limit of a sum like the one which occurs in (13.5) makes its appearance in all problems of the integral calculus mentioned on page 350, and in many others. Their general character is as follows: A function $f(x)$ of the independent variable x is given for values of x on the interval (ab); this interval is divided into n parts. These parts may be equal (as they are in the example which we are discussing), but are not necessarily equal. The length of each part is multiplied by the value of this function at some point within the part or at one of its endpoints. These products are then added together. We consider then the limit of this sum as n, the number of parts, is increased indefinitely in such a way that the length of each interval becomes an infinitesimal. In notational representation we have the following form:

Let the values of the independent variable at the points by means of which the interval (ab) is divided be

$$x_0 = a,\ x_1,\ x_2,\ \ldots\ x_{n-1},\ x_n = b \text{ (see Fig. 89)};$$

FIG. 89

and let the points chosen in these parts correspond to the values

$$X_1,\ X_2,\ \ldots\ X_{n-1},\ X_n$$

of the independent variable. Then we have to consider

$$\lim_{n \to \infty} [(x_1 - x_0)f(X_1) + (x_2 - x_1)f(X_2) + \cdots + (x_n - x_{n-1})f(X_n)]$$

$$= \lim_{n \to \infty} \sum_{k=1}^{n} (x_k - x_{k-1})f(X_k).$$

If this limit exists and if its value is the same for all choices of the numbers $x_1, x_2, \ldots x_n, X_1, X_2, \ldots X_n$ and for all ways in

[1] The symbol $n \to \infty$ is read "as n increases indefinitely."

which the lengths of the parts can tend to zero as n is increased indefinitely, this limit is called the *definite integral from a to b of the function $f(x)$*. In brief, and speaking rather inaccurately, we have:

Definition LII. The definite integral from a to b of the function $f(x)$ defined for values of x from $x = a$ to $x = b$ is the limit of the sum

$$\sum_{k=1}^{n} (x_k - x_{k-1})f(X_k),$$

as $n \longrightarrow \infty$.

The notation used for this definite integral is $\int_a^b f(x)dx$.

The problem with which we are concerned has led us in (13.5) to a limit of exactly this type. The independent variable is called y, the function which is involved is $\dfrac{\delta d}{(d^2 + y^2)^{\frac{3}{2}}}$, the interval concerned is that from $y = 0$ to $y = l$, it has been divided into equal parts, and the point for each part at which the value of the function is taken is the right-hand endpoint. All these particular choices are clearly included in the general definition. We can therefore put (13.5) in the following form:

$$| F | = 2\int_0^l \frac{\delta d}{(d^2 + y^2)^{\frac{3}{2}}} dy.$$

The question which now clamors for an answer is: What has this definite integral from a to b as defined in Definition LII to do with the indefinite integral between the limits a and b as discussed on page 345? The answer is very simple to state, viz: *the two are equal;* but it is not so simple to prove — we shall come back to this point presently. If we accept the answer we can complete the solution of our problem by determining a function of y whose derivative is equal to $\dfrac{\delta d}{(d^2 + y^2)^{\frac{3}{2}}}$. Since our integration-technique is not quite equal to such a task let us simply introduce the notation $\phi(y)$ to designate such an indefinite integral of $\dfrac{\delta d}{(d^2 + y^2)^{\frac{3}{2}}}$; then the magnitude of the force which the bar AB exerts on the unit mass at P is equal to $2[\phi(l) - \phi(0)]$. This completes our study of the particular problem, except for two left-overs, one very special,

the other of general scope, viz.: (1) the determination of a function $\phi(y)$ whose derivative is $\frac{\delta d}{(d^2 + y^2)^{\frac{3}{2}}}$; (2) the proof of the equality of the definite integral from a to b and the indefinite integral between the limits a and b.

The first of these is a matter of technique — it takes some time and patience. The reader who has learned the use of the few tools with which we have been working, will be able to verify by means of *138*, 24 and 26, and *145*, 16, that $\frac{\delta y}{d(d^2 + y^2)^{\frac{1}{2}}}$ is such a function. From this we obtain as a final result that the magnitude of the required attraction is equal to $\frac{2\delta l}{d(d^2 + l^2)^{\frac{1}{2}}}$; this can be written in the form $\frac{M}{da}$, where M is the mass of the bar, d the perpendicular distance from P to the bar, and a the distance from P to either end of the bar (compare Fig. 87).

The second point must be reserved for a fuller discussion (see *149*). The detailed treatment which we have given of the attraction of the bar AB upon a unit mass at P has found its justification in the fact that it has acquainted us with some of the methods and fundamental principles of the integral calculus.

149. An important link. We have stated without proof that the definite integral from a to b of a function $f(x)$ is equal to the indefinite integral of $f(x)$ between the limits a and b. In this statement we have what is usually called the "fundamental theorem of the integral calculus." A proof of this theorem, which makes possible, as a deus ex machina, the final solution of the attraction problem, as well as of many other similar problems, is not within our power at present. But we have learned enough to be able to understand why it is at least plausible.

Let us follow geometrically, step by step, the presentation on page 356. If the function $f(x)$ has the curve represented in Fig. 90, and the division of the interval (ab) is made in a manner similar to that illustrated in Fig. 89, then the products $(x_1 - x_0)f(X_1)$, $(x_2 - x_1)f(X_2)$, . . ., represent the areas of the rectangles $Q_0x_0x_1P_1$, $Q_1x_1x_2P_2$, $Q_2x_2x_3P_3$, . . . $Q_{n-1}x_{n-1}x_nP_n$.

Consequently the sum whose limit is defined as the definite integral of $f(x)$ from a to b is connected in an intimate way with the

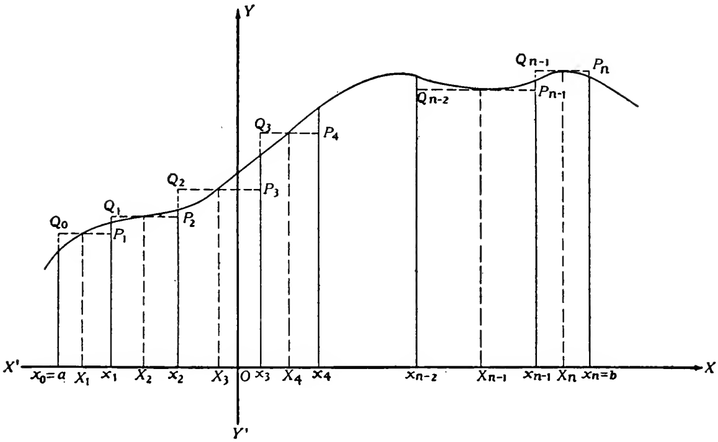


FIG. 90

area bounded by the curve, the X-axis and the ordinates $x = a$ and $x = b$. But it is exactly that area which we have previously obtained from the indefinite integral of $f(x)$ between the limits a and b. These few hints should be sufficient to make the fundamental theorem of the integral calculus acceptable to the reader; for a more satisfactory treatment he will by this time know where to go.

On account of the equality of the "definite integral" and the "indefinite integral between limits," the notation commonly used for the latter is identical with that for the former. We have made a distinction between them by omitting dx from the indefinite integral between limits, so that the fundamental theorem can be expressed by the equation

$$\int_a^b f(x)\, dx = \int_a^b f(x).$$

Let us illustrate the theorem by a simple example, viz. by showing that $\int_0^1 x\, dx = \int_0^1 x$. To determine $\int_0^1 x\, dx$, we divide the interval (0,1) in n equal parts, so that each part has length $\frac{1}{n}$; and we multiply the length of each part by the value of the function x at

its right endpoint, which happens to be equal to the abscissa of that endpoint; for the kth part it is $\frac{k}{n}\cdot$ Hence we have

$$\int_0^1 x\,dx = \lim_{n\to\infty}\left[\frac{1}{n}\cdot\frac{1}{n}+\frac{1}{n}\cdot\frac{2}{n}+\frac{1}{n}\cdot\frac{3}{n}+\cdots+\frac{1}{n}\cdot\frac{n-1}{n}+\frac{1}{n}\cdot\frac{n}{n}\right]$$

$$= \lim_{n\to\infty}\frac{1}{n^2}[1+2+3+\cdots+(n-1)+n]$$

$$= \lim_{n\to\infty}\frac{1}{n^2}\cdot\frac{n(n+1)}{2}{}^{1} = \lim_{n\to\infty}\frac{n+1}{2n}$$

$$= \lim_{n\to\infty}\left(\frac{1}{2}+\frac{1}{2n}\right) = \frac{1}{2}.$$

On the other hand, for $\int_0^1 x$, we need to find a function whose derivative equals x; we know that $\frac{x^2}{2}$ is such a function. The difference between its values at 0 and 1 is equal to $\frac{1}{2}-\frac{0}{2}=\frac{1}{2}$. The fundamental theorem is thus verified.

It is easy to carry out the calculation of the definite integral in this example because the sum $\Sigma\,(x_k - x_{k-1})f(X_k)$ is an arithmetical progression. As soon as the function $f(x)$ becomes more complicated, the corresponding sum can not be determined very easily. It is from this fact that the fundamental theorem of the integral calculus derives its importance. For it makes possible the calculation of definite integrals by means of indefinite integrals between limits. Its effectiveness depends of course upon the ability to find indefinite integrals.

150. Little drops of water make the mighty ocean.[2]

1. Use the methods of the integral calculus to determine the volume of a right circular cone of height h and base-radius R.

2. Determine the volume of a sphere of radius R by the use of integral calculus.

[1] Compare the footnote on p. 24.

[2] The reader is likely to experience difficulties in trying to answer some of the questions in this section (for instance in 6, 8, 9, 11; less in 1, 2, and 12), chiefly because it is not always easy to discover what "drops" to choose to construct the "ocean" — in technical language, to find the most suitable "element of the definite integral" (compare *151*). Perhaps these questions will provide points of interest to return to at a later time.

3. Show that the definite integral of x^2 from 0 to 1 is equal to the indefinite integral between the limits 0 and 1. (It will be useful to have available the information that the sum of the squares of the first n integers, i.e. $1^2 + 2^2 + 3^2 + \cdots + n^2$ is equal to $\frac{n(n+1)(2n+1)}{6}$.)

4. Determine the area enclosed by the curve $y = \frac{1}{x}$, the X-axis and the ordinates $x = 1$ and $x = 2$. (Make use of *144*, IV and obtain an approximate value for the natural logarithm of 2.)

5. Use the fundamental theorem of the integral calculus to find the definite integral of:

(a) the function $\sqrt{x+1}$ from $x = 0$ to $x = 7$;

(b) the function $\frac{1}{(x+1)^2}$ from $x = 0$ to $x = 3$;

(c) the function $(x^2 + 3)^4 x$ from $x = 0$ to $x = 5$.

(Valuable assistance can be furnished here by *145*, 10 and 16.)

6. Determine the attraction of a bar of length l upon a unit mass at a point on the extension of the bar and at a distance d from one end of the bar.

7. A point moves on a straight line in such a way that its acceleration at time t sec. is equal to $2t - 3$; it started from rest at time $t = 0$. Determine its velocity and its distance from the starting point 5 seconds afterwards.

8. A circular cylinder of variable density is 10 ft. high and 2 ft. in diameter. The density along its axis is $\frac{1}{2}$; at a distance r ft. from the axis the density is equal to $\frac{r}{5} + \frac{1}{2}$. Determine the mass of this cylinder.

9. Water flows into a tank at varying speed. At the start, the speed is 5 gallons per minute; after t minutes the speed is $\frac{t}{2} + 5$ gallons per minute. Determine how much water there is in the tank after 15 minutes.

10. The inner radius of a flat circular ring of thickness d is 5 inches, the outer radius 10 in.; it is of uniform density δ. Determine the attraction upon a unit mass at the center.

11. Solve the corresponding problem for one half of this ring.

12. A wedge has the form of a right triangular prism (see Fig. 91). The dimensions of the rectangular base $ABCD$ are 15 in. by 4 in.; the face PAB is an isosceles triangle whose base AB is 4 in. in length, and whose height is 6 in. The density at the base is 3, and the density at height h above the base is $3 - \frac{h}{3}$. Determine the mass of the wedge.

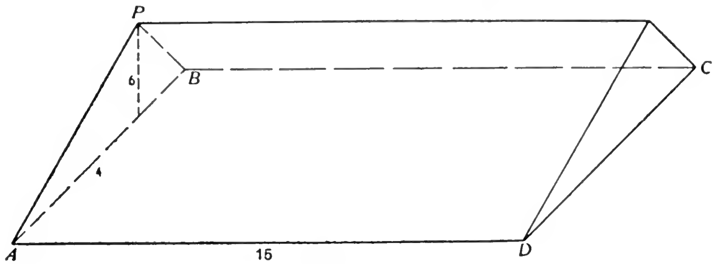


FIG. 91

151. Greater worlds to conquer. The terms of a sum whose limit is a definite integral, as for instance the terms $\dfrac{\delta d(y_k - y_{k-1})}{(d^2 + y_k^2)^{\frac{3}{2}}}$ in (13.5) or the term $(x_k - x_{k-1}) \cdot f(X_k)$ in definition LII, tend to zero as the number n of the parts, into which the interval of the independent variable is divided, increases indefinitely. These terms are therefore infinitesimals; their number increases indefinitely as each term tends to zero. Speaking roughly, we can say that a definite integral is the sum of an *infinite number of infinitesimals.* The general term of such a sum is frequently called the *element of the definite integral.* It is indeed the central problem of the integral calculus to determine such a sum of an infinite number of infinitesimals, just as it is the central problem of the differential calculus to determine the limit of the ratio of two infinitesimals (compare *140*). In the examples which have been treated and in the questions which have been proposed to the reader, the sums consist of a *single infinitude* of terms, i.e. the elements of the definite integrals occurring in them depend on a single variable subscript k, they form a set with one degree of freedom (compare *131*). In a problem for whose solution a definite integral has to be set up, the essential difficulty frequently consists in selecting the most convenient element (compare footnote 2 on p. 360).

In more advanced parts of the integral calculus we have to deal with limits of sums of infinitesimals which depend on two, three or more independent variable subscripts, i.e. with double, triple infinitudes of infinitesimals, or with infinitudes of still higher orders.[1] The limit of the sum of a double (triple) infinitude of

[1] The reader will recall that he has met double and triple infinitudes in Chapter II, where he became acquainted with the set of points in the plane with integral coördinates

infinitesimals is called a *double* (*triple*) *integral;* in general we speak of *multiple integrals.* A few examples of such integrals will aid in acquainting us with the spirit of the integral calculus.

1. To determine the area A, enclosed by a closed plane curve C, we draw a set of lines parallel to the X-axis and a set of lines parallel to the Y-axis, a distance e apart (see Fig. 92) so as to cover

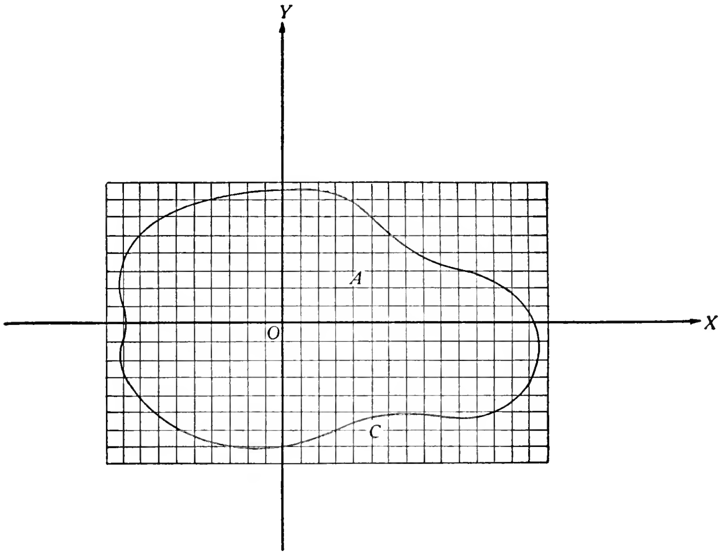


FIG. 92

the region enclosed by the curve with a network of square cells. If we denote by m the sum of the areas of those cells which are entirely contained within C, and by M the sum of the areas of the cells needed to enclose the area A, bounded by C, we can say that

$$m \leqq A \leqq M. \tag{13.6}$$

Next it is shown that, as the network is made finer and finer, both M and m tend towards a limit, so that if the difference $M - m$ tends to zero, as $e \to 0$, M and m will tend towards the same limit

(see pp. 22, 23) and the set of ordered triads of natural numbers (see p. 26). At that time we were concerned only with the sets as a whole; now we are interested in the "*sums*" of the elements of such sets.

(Theorem LXXVI). Whether $M - m$ does tend to a limit depends upon the character of the curve C and may not be easy to decide. If it does, we can use the fly-swatter to show that the common limit is the desired area A. But the cells which make up m or M have two degrees of freedom, viz. the abscissa *and* the ordinate of one of their corner points. Thus it is seen that A is the limit of a double infinitude of infinitesimals; it is therefore represented by a double integral. How to calculate such an integral, that is still a different story.

2. Let us consider a solid V, like a sphere or a cube, whose density varies from point to point in a known manner; whose density, in other words, is a known function say $\delta(x, y, z)$ of the coördinates of a point in the region enclosed by the solid. To determine its mass, we construct three sets of parallel planes, any two planes of different sets being mutually perpendicular, so as to divide the region of space in which the solid lies in cubical cells. If the actual mass of the solid is M, and if we denote the edges of the cubical cells by $x_{i+1} - x_i$, $y_{j+1} - y_j$ and $z_{k+1} - z_k$, we can set up an inequality similar to (13.6). For, if $\underline{\delta}(x_i, y_j, z_k)$ is the maximum density in the cell of which one corner has the coördinates (x_i, y_j, z_k) and $\underline{\delta}_1(x_i, y_j, z_k)$ the minimum density in this cell, we find

$$\begin{aligned}\Sigma\underline{\delta}_1(x_i, y_j, z_k)(x_{i+1} - x_i)(y_{j+1} - y_j)(z_{k+1} - z_k) \leqq M \\ \leqq \Sigma\underline{\delta}(x_i, y_j, z_k)(x_{i+1} - x_i)(y_{j+1} - y_j)(z_{k+1} - z_k).\end{aligned}$$

The sum on the left covers all the cells which are entirely within the solid; that on the right all those cells necessary to enclose the solid. After it has been shown that the difference between these two sums is an infinitesimal, when the edges of the cells tend to zero, it follows, by an argument which is now familiar, that M is equal to the limit of either of the sums, provided such limits exist.

Each of the sums consists now of terms which depend on three variable subscripts; hence M becomes a triple integral. It is represented by $\iiint_V \delta(x, y, z)\, dx\, dy\, dz$. This process is frequently described by saying that the mass of a solid is determined by "integrating the density over the region of space occupied by the solid."

Two further examples of integration will bring this chapter to a close.

3. To determine the mass of a curved wire C of known variable density, we integrate the density over the wire. If the distance along the wire from a fixed point A to a variable point P is denoted by s the density will be a known function of s say $f(s)$. The element of the definite integral will then be $(s_{k+1} - s_k)f(s_k)$ and the mass of the wire is represented by $\int_C f(s)ds$. This kind of integral is called a *line-integral.*

4. If we wish to determine the mass of a thin curved shell of known variable density, we proceed in a similar manner; an example would be furnished by the metal covering of a roof. The mass will be obtained by integrating the density over the surface. The analytical representation is more complicated in this case, but the essential features of the method are the same as those which we have used before; the result will be a *surface integral.*

In the discussion of these examples we have confined ourselves to an indication of the general procedure. To carry out the details and to obtain means of calculating such integrals a good many things have to be taken account of which we have left out of consideration. Moreover we have been limited to a restricted type of problem. In various domains of science, especially in physics, questions of the same general character arise. Any one who has thoroughly grasped the fundamental aspects of the integral calculus will be able to apply them to such problems after he has acquired an understanding of their particular character. To bring them to the final state of solution he would moreover have to acquire the technique of the subject. This is not our game.

CHAPTER XIV

AN APPROACH TO THE SECRETS OF NATURE

We may doubt the warranty of the priest, but never that of the mathematician. And the successful launch of a ship is the final solution to a host of converging problems. — H. W. Tomlinson, *All Our Yesterdays*.

152. The object in view. In the present chapter we hope to come a little closer than has been possible so far to the way in which mathematical analysis deals with phenomena of the physical world. It would not be difficult for one who has acquired a somewhat greater amount of technical knowledge to get insight into the methods which are used for this purpose. In the absence of such knowledge we shall have to content ourselves with the study of a few simple special cases. It will be our aim, as in the preceding chapters, to deduce from these particular examples enough of the general principles to give an idea of the way in which more complicated problems can be dealt with.

153. Some reconnoitring steps. We start with some very simple problems, similar to some which we have discussed in the preceding chapter, viz. the following:

1. A point moves along a straight line in such a way that its velocity (measured in feet per second) at any time t (measured in seconds) is equal to 3 times t. What will be its distance at $t = 5$ from its position at $t = 0$? Let O (Fig. 93) be its position at $t = 0$,

FIG. 93

and let us denote by x the distance from O to the point P which is reached at time t. Our problem consists then in determining x as a function of t. The velocity of the moving point is determined by the derivative with respect to t, i.e. by $D_t x$ (compare p. 329). Consequently, the analytical formulation is as follows:

To determine x as a function of t, in such a way that

$$(14.1) \quad (1)\ D_t x = 3t, \quad \text{while } (2)\ x = 0, \quad \text{when } t = 0.$$

The first of these conditions is satisfied by any indefinite integral of $3t$. Consequently the function we are seeking must be contained among the functions which are represented by

$$x = \frac{3t^2}{2} + \text{constant}$$

(compare Theorem LXXIV, and *144*, III). It remains to determine the constant so as to comply with condition (2) in (14.1). It follows at once that this constant must be 0, so that the distance of the moving point from O at time t is equal to $\frac{3t^2}{2}$; hence after 5 seconds the distance will be $37\frac{1}{2}$ feet.

2. Let us next consider a point which moves along a straight line in such a way that its acceleration (measured in feet per sec., per sec.) at any time t (measured in seconds) is equal to 3 times t. If we know moreover that its velocity at time $t = 0$ is 5 ft. per sec., we wish to determine its distance from the starting point at time $t = 4$.

If we use the notations of the preceding problem, the question reduces to that of finding x as a function of t, such that its second derivative is equal to $3t$, while at $t = 0$ the first derivative is equal to 5 and the function itself equal to 0. The analytical formulation is therefore the following:

To determine x as a function of t such that

(14.2) (1) $D_{tt}x = 3t$, (2) at $t = 0$, $D_t x = 5$ and $x = 0$.

The solution proceeds in two steps: First we observe that the first derivative, $D_t x$, must be an indefinite integral of $3t$ and therefore one of the functions represented by $\frac{3t^2}{2}$ + constant; but since for $t = 0$, $D_t x = 5$, we conclude that the constant must be 5 so that we have $D_t x = \frac{3t^2}{2} + 5$. In the second step, we determine x as an indefinite integral of $\frac{3t^2}{2} + 5$. If we recall *144*, III once more we obtain

$$x = \frac{t^3}{2} + 5t + \text{constant};$$

the condition $x = 0$ when $t = 0$ shows that the constant is 0, so that the distance from O at time t is $\frac{t^3}{2} + 5t$, and therefore equal to 52 ft. at $t = 4$.

3. Finally we take a geometrical problem: To determine a curve which passes through the point (1, 1) and on which the abscissa of any point P is equal to one half the distance from the point Q where the line tangent to the curve at P meets the X-axis to the projection R of P on the X-axis. (See Fig. 94.)

If the curve has the equation $y = f(x)$, and P is any of its points, the equation of the tangent line at P is $Y - y = y'(X - x)$, X

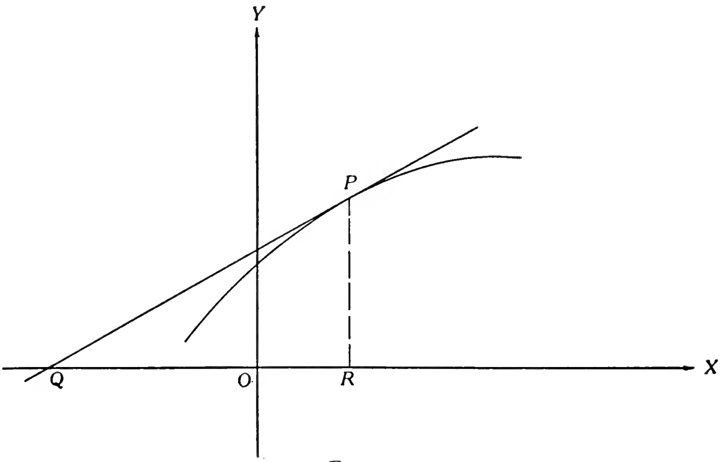


FIG. 94

and Y being coördinates of a variable point on the tangent line (compare p. 317). The distance OQ is the X-coördinate of the point on this line for which the Y-coördinate equals 0; putting 0 for Y in the equation of the tangent line and solving for X we find $OQ = x - \frac{y}{y'}$. The condition of the problem, apart from the requirement that the curve pass through the point (1, 1) is that $OR = \frac{QR}{2}$, i.e. that $QO = OR$; this analytical formulation leads to the equation

$$\frac{y}{y'} - x = x.$$

A simple transformation shows now that the problem asks for the determination of y as a function of x such that

$$(14.3)\quad (1)\ 2x \cdot D_x y = y \quad \text{and} \quad (2)\ y = 1 \text{ when } x = 1.$$

This is a somewhat different question from those that occurred in the preceding examples; the reader may well be at a loss to know how to solve it. A little technique would be very useful here; in its absence, a good memory would be of service. If he recalls *138*, 9 or *144*, VI, he will know that if $y = \sqrt{x}$, then $D_x y$ is equal to $\frac{1}{2\sqrt{x}}$, so that $2x \cdot D_x y = \frac{2x}{2\sqrt{x}} = \sqrt{x} = y$. Knowing this, he can readily show that any function of the form $c\sqrt{x}$, in which c is a constant, satisfies condition (1) in (14.3).[1] If we are willing to assume that the difficulty referred to in the footnote has been overcome, it remains to determine the constant c in such a way that condition (2) in (14.3) is also satisfied. This clearly leads to the solution $c = 1$, so that $y = \sqrt{x}$ is the equation of the required curve. The curve can be obtained by plotting points whose coördinates satisfy this equation.

154. An outlook. The problems treated in the preceding section have been given analytical formulations in the equations and conditions (14.1), (14.2) and (14.3). The three equations are evidently of a character quite different from that of the equations to which the so-called "word problems" of algebra lead. They are like them to the extent that they are also equations, i.e., questions. But, whereas the algebraic equations ask for values of one or more variables, these equations require the *determination of a function*, for which there exists a given relation between the independent variable, the function and a derivative of this function; the functions must moreover take a given value for a particular value of the independent variable. Equations of this kind are called *differential equations*. The general definition is as follows:

Definition LIII. A differential equation is a question which calls for the determination of a function which satisfies a specified relation into which may enter the independent variable, the function, one or more of its derivatives and any number of constants.

[1] But even so, he would not yet know that every function which satisfies (1) must be of the form $c\sqrt{x}$.

In equation (14.1) occur only the independent variable, t, the first derivative of a function of this variable and a constant; in equation (14.2) the independent variable, the second derivative of a function of this variable and a constant, in equation (14.3), the independent variable, x, a function of x, its first derivative and a constant. Equations (14.1) and (14.3) are said to be of the *first order;* equation (14.2) is one of the *second order.*

The special condition which the required function is expected to satisfy is called the *initial condition*, or the *boundary condition.*

It is a fact of outstanding significance that many important problems in various domains of science find their analytical formulation in *differential equations with initial conditions.* A study of such problems leads to differential equations whose solution requires a good deal of technical knowledge of mathematics. We can not expect to go very far on the basis of such preparation as we have acquired. Indeed, the examples of differential equations which we have taken up so far give a totally inadequate idea of the subject; they are too elementary in character to bring out its important features. For these reasons we must become acquainted with some additional facts from the differential and integral calculus. After that we can go a little more deeply into the domain of differential equation, so as to get a better idea of its scope and its fundamental character. But it will be well to preface this further study by making sure that the first simple ideas have been mastered.

155. Getting a foothold.

1. A point moves along a directed straight line with an origin O, starting at time $t = 0$ from A, a distance of 3 units to the left of O. Its velocity at time t is equal to $2t - 1$. Determine its position at time $t = 4$.

2. Determine a curve through the point $(1, -1)$ such that the tangent line to the curve at any point P meets the Y-axis in a point Q such that the distance from Q to the projection P' of P on the Y-axis is equal to 3 times the ordinate of P.

3. A point moves along the X-axis with uniform acceleration of a ft. per sec., per sec. It starts from a point A of abscissa 5 with a velocity of 10 ft. per sec. Determine its position and its velocity at time $t = 8$.

4. Solve a problem similar to 3, but with initial position x_0 and initial velocity v_0; determine its position and velocity at time t.

5. Determine a curve through the point $(1, 4)$ such that the tangent line at any point P meets the X-axis in a point Q whose distance to the projection of P on the X-axis is equal to 3 times the abscissa of P.

6. The acceleration at time t of a point moving along the X-axis is equal to at. It starts at time $t = 0$ from the point of abscissa x_0 with a velocity equal to v_0. Determine its position and velocity at time t.

156. More supplies are needed. It will have become clear that the knowledge of a great many functions and of their properties will increase the likelihood of being able to find the answer to a differential equation. We shall therefore extend somewhat our acquaintance with functions and call attention to new properties of some functions which have been met before.

In Chapter VI we encountered for the first time the number e (see pp. 120–123), the base of the natural logarithms. In *144*, IV, it

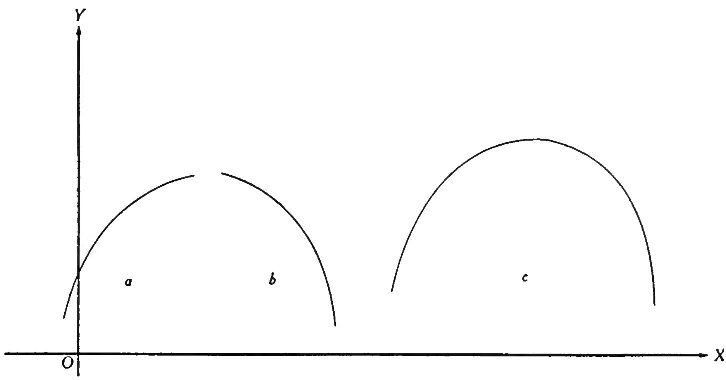


FIG. 95

was stated that $D_x \log_e x = \frac{1}{x}$, i.e. that $\log_e x$ is an indefinite integral of $\frac{1}{x}$ (see p. 346). From this latter fact we are going to derive an important consequence.

Let us look once more at Definition XLVIII, in which the concept "function" was introduced. When "the two variables x and y are so related that to each value of x chosen from the whole or a part of the set C of real numbers, there corresponds a single real value of y," it may very well happen that, at the same time to each value of y chosen from the whole or from a part of the set C, there corresponds a single real value of x. In other words, when y is a function of x, it may happen that at the same time x is a function

of y. This would clearly be the case if the curve corresponding to the function is of the type indicated in Fig. 95, a and b, but not if it has the character of c. This consideration leads us to the following definition:

Definition LIV. When the two variables x and y are so related that they determine y as a function of x *and also* x as a function of y, these two functions constitute a *pair of inverse functions.*

For example, the relation $3x + 5y = 4$ gives rise to $y = \dfrac{4 - 3x}{5}$ expressing y as a function of x and also to $x = \dfrac{4 - 5y}{3}$ which expresses x as a function of y. These functions, i.e. the ways in which the dependent variable (y in the first case, x in the second) depends on the independent variable (x in the first case, y in the second), constitute a pair of inverse functions. The emphasis is placed most readily on the functional dependence if in both cases the dependent variable is denoted by one and the same letter, say v, and the independent variable by another letter, say u. The two inverse functions are then represented as $v = \dfrac{4 - 3u}{5}$ and $v = \dfrac{4 - 5u}{3}$.

We remember (see Definition XXVI) that if $b = a^c$, then $c = \log_a b$, and conversely. Hence if $y = e^x$, then $x = \log_e y$; this means that $v = e^u$ and $v = \log_e u$ is a pair of inverse functions.

An interesting example is furnished by the circle. We have seen (compare p. 303) that the upper half of the unit-circle with center at the origin is represented by $y = \sqrt{1 - x^2}$. This curve is of the type shown in Fig. 95, c. But if we limit x to positive values between 0 and 1, i.e. if we take $0 \leqq x \leqq 1$, we obtain an arc of the type b. It is easily seen that the relation between x and y which holds on this arc also gives rise to $x = \sqrt{1 - y^2}$, provided y is in its turn limited to the range (0, 1). The two inverse functions are then $v = \sqrt{1 - u^2}$ and $v = \sqrt{1 - u^2}$, i.e. they are identical; in other words the function $v = \sqrt{1 - u^2}$, $0 \leqq u \leqq 1$ is its own inverse. A relation between x and y which gives rise to two identical functions is called *symmetric in x and y.* Another example of a symmetric relation is given by the equation $x + y = c$, which gives rise to

the same function, whether it is solved for y in terms of x or for x in terms of y; see also *160*.

We inquire now as to the relation between the derivatives of the two functions which are inverses of each other. For this purpose we follow carefully Definition XLIX and the discussion preceding it (see *135*). Suppose that $v = f(u)$ and $v = F(u)$ is a pair of inverse functions. If $u = a$, $v = b$ are values of the variables which satisfy the first of these equations, i.e. if $b = f(a)$, then $u = b$, $v = a$ are values which satisfy the second, i.e. $a = F(b)$. Similarly, to the pair of values $(a + h_1, b + k_1)$ belonging to the first equation corresponds the pair $(b + k_1, a + h_1)$ of the second; similarly for $(a + h_2, b + k_2)$ and $(b + k_2, a + h_2)$, etc. Consequently, since the derivative of a function is the limit of the ratio of the change in the dependent variable to that in the independent variable, we conclude that if $f'(a) = \lim_{h \to 0} \frac{k}{h}$, then $F'(b) = \lim_{k \to 0} \frac{h}{k}$. But we already know that $k \to 0$ as $h \to 0$ (compare *140*), so that we may also write $F'(b) = \lim_{h \to 0} \frac{h}{k}$. There remains to be settled the question whether $\frac{h}{k}$ tends to a limit as $h \to 0$, if it is known that $\frac{k}{h}$ tends to a limit. Let us suppose then that as $h \to 0$, the ratio $\frac{k}{h}$ tends to L and let us prove the following useful theorem:

Theorem LXXVII. If the variable x approaches the limit L, then the variable $\frac{1}{x}$ tends to the limit $\frac{1}{L}$ provided $L \neq 0$.

Proof. In the terminology and notation which we have frequently used,[1] the hypotheses of the theorem are that $|L - x| < \epsilon$ and $L \neq 0$, and the conclusion is that $\left|\frac{1}{L} - \frac{1}{x}\right| < \epsilon$. Now $\frac{1}{L} - \frac{1}{x} = \frac{x - L}{Lx}$, so that

$$\left|\frac{1}{L} - \frac{1}{x}\right| = \frac{|L - x|}{|Lx|}. \tag{14.4}$$

[1] On account of the close proximity to the appearance of e as the base of the natural logarithms, we use here ϵ instead of e to designate the "arbitrarily preassigned positive number no matter how small"; compare footnotes 1 on p. 296, 1 on p. 339 and 2 on p. 354.

From the hypotheses it follows that the variation of x will ultimately surely be outside the interval from o to $\frac{L}{2}$, i.e. $|x| > \frac{|L|}{2}$, and hence $\frac{1}{|x|} < \frac{2}{|L|}$. It follows therefore from (14.4) that

$$\left|\frac{1}{L} - \frac{1}{x}\right| < \frac{2\epsilon}{L^2}.$$

Since L is a fixed number and ϵ arbitrarily small, we conclude that $\left|\frac{1}{L} - \frac{1}{x}\right|$ can be made as small as may be wished by taking $|L - x|$ small enough. This means that, when $L \neq 0$ and $L - x$ is an infinitesimal, then $\frac{1}{L} - \frac{1}{x}$ is also an infinitesimal, as was to be proved.

This theorem finds immediate application in the problem we are considering. It enables us to answer in the following way the question to which we were led: If $f'(a) \neq 0$, then $F'(b) = \frac{1}{f'(a)}$, where $a = F(b)$. Since b is an arbitrary value of the independent variable u in $F(u)$, we have obtained a solution of our problem, viz..

Theorem LXXVIII. If $v = f(u)$ and $v = F(u)$ are a pair of inverse functions, then $F'(u) = \frac{1}{f'[F(u)]}$, for all values of u for which $f'[F(u)] \neq 0$.

Let us take $f(x) = \log_e x$; then $f'(x) = \frac{1}{x} \neq 0$ for all values of x. The inverse function of $\log_e x$ is e^x; i.e. $F(x) = e^x$. Hence

$$F'(x) = \frac{1}{f'[F(x)]} = \frac{1}{\frac{1}{e^x}} = e^x.$$

Thus we have obtained a valuable addition to our knowledge of derivatives, viz.:

Theorem LXXIX. The derivative of the function e^x is equal to e^x.

Corollary. If c is a constant, the derivative of the function e^{cx} is equal to ce^{cx}.

The proof of this corollary is also a consequence of Theorem LXXVIII (see *160*, 10).

The corollary expresses an important property of the exponential function e^{cx}, for it says that the rate of change of this function is proportional to its value. It means, if $c > 0$, that the larger the value of the function the more rapidly it will increase for a fixed increase in x; if $c < 0$, the derivative is negative, so that the rate of decrease of the value of the function for increasing x is proportional to its size. It is this property of the function e^{cx} which makes it an important aid in the study of organic growth, for in this process the rate of change is often approximately proportional to the size. Since the nourishment of a tree depends at least in part on the chlorophyl content of the leaves, its rate of growth is related to the total area of its leaves. The rate of decay of an organic mass depends upon the amount of undecayed matter that is exposed to contact with air; the process will be most rapid at the start and will slow up as the decay advances. Other illustrations are found in the theory of compound interest, in bacterial growth and in many other fields.

157. New tools are made. With the aid of Theorems LXXVIII and LXXIX, we can extend our knowledge of the differential calculus a good deal. We shall bring the corollary of the latter theorem in connection with the formula of Euler,

$$e^{ix} = \cos x + i \sin x, \tag{6.01}$$

which was mentioned and used, but not proved, in Chapter VI (see p. 120).[1] On the one hand, we obtain then the fact that

$$D_x(\cos x + i \sin x) = D_x e^{ix} = ie^{ix} = i(\cos x + i \sin x) = -\sin x + i \cos x.^{2}$$

On the other hand, we know from *138*, 22 and 24 that

$$D_x(\cos x + i \sin x) = D_x \cos x + iD_x \sin x.$$

Comparison of these two expressions for $D_x(\cos x + i \sin x)$ leads to the desired result,[3] viz.:

Theorem LXXX. The derivative of $\cos x$ is equal to $-\sin x$; the derivative of $\sin x$ is equal to $\cos x$.

Corollary. The derivative of $\sin cx$ is equal to $c \cos cx$; the derivative of $\cos cx$ is $-c \sin cx$.

[1] It may not be superfluous to remind the reader of the fact that this formula presupposes that x is the measure of the angle *in radians* (compare p. 121).

[2] Compare the discussion of the symbol i on pp. 84–85.

[3] Remember also Definition XVI and the discussion in *46*.

This corollary is proved exactly like the theorem itself, starting from the equation obtained from (6.01) by writing cx in place of x (see *160*, 11).

Upon these results we can build further; from the formulas

$$D_x \cos x = -\sin x \quad \text{and} \quad D_x \sin x = \cos x,$$

we obtain by repeated application:

$$\begin{aligned} D_{xx} \cos x &= -\cos x, & D_{xx} \sin x &= -\sin x; \\ D_{xxx} \cos x &= \sin x, & D_{xxx} \sin x &= -\cos x; \\ D_{xxxx} \cos x &= \cos x, & D_{xxxx} \sin x &= \sin x. \end{aligned}$$

If the sequence of derivatives of $\cos x$ and $\sin x$ is continued beyond the fourth, we shall obtain an unending repetition of $-\sin x$, $-\cos x$, $\sin x$, $\cos x$ for the one, and of $\cos x$, $-\sin x$, $-\cos x$, $\sin x$ for the other.

There is a great temptation at this point to continue the development of this rather fascinating chapter of the differential calculus. The reader may yield to the temptation by indulging in the study of one of the many books on this subject; the author must not forget his main purpose however — he must steer his course unaffected by the call of the sirens. The insertion of this section of the differential calculus is unavoidable if we are to secure further insight into the significance of differential equations. To this subject we must now return. Let us record the fact that the results which were obtained in *156* and *157* were attainable for us by assuming without proof that $D_x \log_e x = \dfrac{1}{x}$, and by accepting without proof Euler's formula (6.01).

158. The tools are put to use. We are now ready to consider a few more interesting problems which lead to differential equations.

1. A point moves along a straight line in such a way that its velocity is proportional to its distance from the starting point. To determine its position at any time we have to express its distance, x, from the starting point O (see Fig. 96) as a function of t. We readily see that the analytical formulation of our problem leads to the following differential equation and boundary condition:

(14.5) $\quad$ (1) $D_t x = cx$, $\quad$ (2) $x = 0$ when $t = 0$.

The solution of this equation can not be found in the way that was used for (14.1) and (14.2). While these equations asked for

functions whose derivatives with respect to t are given functions of t, we are now asked for a function $x(t)$ whose derivative with respect to t is a known function of x. The question is more nearly like that asked in (14.3). The theory of differential equations has a systematic procedure for its solution. We have to rely here on such knowledge as we have of the differential calculus. Having so recently become acquainted with Theorem LXXIX and its corollary, we recognize without difficulty that $x = e^{ct}$ is a solution of the equation (14.5) and that even $x = ke^{ct}$, in which k is an arbitrary constant, is also a solution. That there are no other solutions — this we can not prove as yet. Among the functions ke^{ct} we have to look for one which satisfies the boundary conditions (2). Since $ke^{co} = k$ the only value of k which will make $x = 0$, when $t = 0$, is $k = 0$. But this gives us the solution $x = 0$ for our problem; i.e. a point which is subject to the conditions stated in this problem would never get away from the starting point. Does this seem reasonable to you, non-technical reader?

2. Suppose now that we modify the problem in 1 by requiring again that the velocity of the point P is proportional to the distance

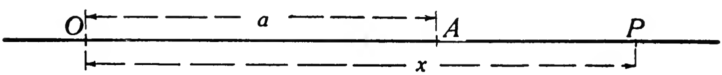


FIG. 96

OP from the point O, and that the starting point is at A, so placed that $OA = a \neq 0$. The analytical formulation of the problem is now:

$$(14.6) \qquad (1)\ D_t x = cx, \quad (2)\ x = a, \quad \text{when } t = 0.$$

Since the change is only in the initial condition, we have again the possibilities contained in the solution $x = ke^{ct}$. The initial condition now requires that $a = ke^{co}$, i.e. $k = a$. The solution of the problem is therefore $x = ae^{ct}$. Since c and a are given constants, this formula enables us to determine the position at any time t. Moreover the velocity at time t is equal to cae^{ct}.

Comparison of examples 1 and 2 brings out the importance of the initial conditions. A mere change in the conditions at birth may seriously affect the course of a life — both the differential equation and the initial conditions are of significance in determining future events.

It is worth while to remark moreover that, if the constant c is positive, the point will move farther and farther away from the point O; if c is negative the point will move towards the origin.

3. If a point moves along a line in such a way that the acceleration is proportional to the distance from the starting point, we have a more interesting situation. Using the notation of Fig. 96, the differential equation becomes

$$D_{tt}x = cx. \tag{14.7}$$

Does our enlarged knowledge of differential calculus enable us to find a solution of this differential equation? Each reader will deal with this question in his own way. But there is no doubt that every one can verify on the strength of our previous work that $x = e^{t\sqrt{c}}$ is indeed a solution of (14.7), and that whether $\sqrt{c}$ designates the positive square root of c or the negative square root. Hence we have at least two solutions, viz. $x = e^{t\sqrt{c}}$ and $x = e^{-t\sqrt{c}}$, where now $\sqrt{c}$ designates, in agreement with our convention, the positive square root of c. What is the most general solution of our equation? For the present we shall answer this question categorically reserving a partial justification for later (compare *162*). The most general solution of equation (14.7) is:

$$x = k_1e^{t\sqrt{c}} + k_2e^{-t\sqrt{c}}, \tag{14.8}$$

where k_1 and k_2 are arbitrary constants. Moreover, the velocity at arbitrary time t for this point is given by

$$v = k_1\sqrt{c}\, e^{t\sqrt{c}} - k_2\sqrt{c}\, e^{-t\sqrt{c}}.$$

The distinction between the cases in which $c > 0$ and those in which $c < 0$ is here of fundamental importance. In the former case there is nothing left to do except to determine the constants k_1 and k_2 in such a way as to satisfy the initial conditions, which must give the position *and* the velocity at some fixed time. The situation in the latter case is best understood by means of the Euler formula (6.01). If $c < 0$, there is a real number a such that $c = -a^2$. The solution (14.8) is then transformed as follows:

$$x = k_1e^{iat} + k_2e^{-iat} = k_1(\cos at + i \sin at) + k_2(\cos at - i \sin at);$$

i.e.

$$x = c_1 \cos at + c_2 \sin at, \tag{14.9}$$

where c_1 and c_2, equal to $k_1 + k_2$ and to $i(k_1 - k_2)$ respectively, are again arbitrary constants. It will be instructive for the reader to verify that (14.9) satisfies the equation (14.7) for arbitrary values of c_1 and c_2.

159. Approach to an "actual" problem. The solution (14.9) which has been obtained for equation (14.7) provides a good example to illustrate the importance of differential equations in the study of natural phenomena, to which allusion was made at the beginning of this chapter. We begin with the statement of a fundamental principle of classical mechanics, enunciated by Newton and known as Newton's *second law of motion:*

When a mass is under the influence of a force it acquires an acceleration in the direction of and proportional to the magnitude of the force; the factor of proportionality is the measure of the mass.

If we represent this measure by m and the acceleration by α, the law can be formulated in the statement Force $= m\alpha$.

On the basis of this principle we consider an elastic body, such as a steel spring, or an elastic band. Suppose that it is fastened at one end, A (see Fig. 97), that its natural length is l and that the free end O is pulled out a distance d to the point B, and then let go.

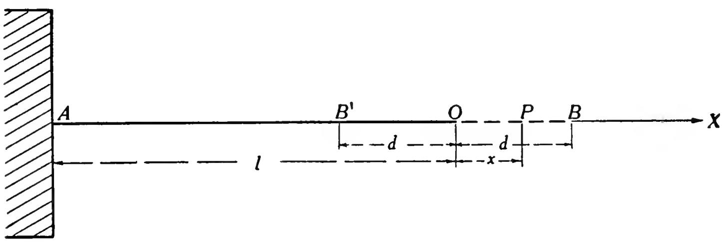


FIG. 97

If we neglect the resistance of the medium in which the body is placed, the free end is subject to only one force, viz. the elastic force. On the basis of experimental evidence it is assumed that this elastic force has a direction opposite to that of the extension and a magnitude proportional to the magnitude of the extension. Hence, in an arbitrary position P, at distance x from O, this force is equal to $-p^2x$; the factor p^2 designates a positive constant, whose magnitude depends upon the nature of the material of which the

body is made. When x is positive, the force is directed negatively; if x is negative it is directed positively — in all cases towards O therefore. We already know that the acceleration of the free end of the body is equal to $D_{tt}x$. Therefore it follows from Newton's second law that the motion of the free end is determined by the differential equation

$$m \cdot D_{tt}x = -p^2x. \tag{14.10}$$

Moreover, the initial conditions are the following:

$$\text{at} \quad t = 0, \quad x = d, \quad \text{and} \quad D_t x = 0. \tag{14.11}$$

If we write a^2 in place of $\frac{p^2}{m}$, the differential equation (14.10) becomes exactly the equation (14.7) for the case in which $c = -a^2$. Hence its general solution is given by (14.9) which represents the distance of the free end from O at any time t. The velocity is obtained from this by use of the corollary to Theorem LXXX; we find:

$$v = -ac_1 \sin at + ac_2 \cos at. \tag{14.12}$$

It remains to determine the constants c_1 and c_2 from the initial conditions (14.11). This is a simple matter, if we remember the contents of *54* (see in particular p. 100). For, we find then from (14.9) that $d = c_1$, and from (14.12) that $0 = ac_2$. The motion of the free end is therefore represented by the equations

$$x = d \cos at \quad \text{and} \quad v = -ad \sin at.$$

Let us pursue the meaning of this result a little further. We know that it was obtained on the supposition that at is measured in radians,[1] so that both x and v are periodic functions of t, and that their values repeat themselves at intervals of 2π in the value of at. Hence both position and velocity of the free end of the body repeat themselves every time a time-period of $\frac{2\pi}{a}$ seconds has elapsed. This means that the free end oscillates back and forth and that the time of an oscillation is $\frac{2\pi}{a}$ seconds. The following table indicates the position and velocity at moments spaced a fourth of the oscillation period apart. In its construction, the values of the sine and

[1] Compare footnote 1 on p. 375.

cosine as given on page 100 are used. After the free end is released at B, it moves to the left with increasing velocity in the negative direction. The velocity reaches a maximum at O; after that the velocity decreases until the point B' is reached. There the return journey begins, with velocity in the positive direction, increasing up to O, and then again decreasing to 0 at the point B. After that the entire process is repeated; and so on indefinitely.

t	at	cos at	sin at	x	Position	v
0	0	1	0	d	B	0
$\frac{\pi}{2a}$	$\frac{\pi}{2}$	0	1	0	O	$-ad$
$\frac{\pi}{a}$	π	-1	0	$-d$	B'	0
$\frac{3\pi}{2a}$	$\frac{3\pi}{2}$	0	-1	0	O	ad
$\frac{2\pi}{a}$	2π	1	0	d	B	0

One further remark. We have seen that the period of oscillation is $\frac{2\pi}{a}$ seconds, and that $a^2 = \frac{p^2}{m}$, in which the denominator is the measure of the mass. Ignoring the significance of p, we see that the larger (smaller) the mass, the smaller (larger) the factor a, and hence the longer (shorter) the period of oscillation. If we know the physical properties which determine the constant p, we can determine their effect on the motion of the free end of the body.

The results we have obtained are in accord with what is suggested by experience. They do not as yet represent any actual experience, because the problem we have treated differs considerably from what occurs in nature. A few exercises are desirable before we approximate more closely to actual conditions (compare *163*).

160. Some independent trips.

1. Determine the inverse function of each of the following functions:

$$\text{(a) } y = 4x + 3; \quad \text{(b) } y = x^3; \quad \text{(c) } y = \frac{x+1}{x-1}.$$

2. Prove that the line joining the points $P(a, b)$ and $Q(b, a)$ is bisected perpendicularly by the line $y = x$, which bisects the angle YOX between the positive halves of the Y- and X-axes.

Remark. In general, if a line l is the perpendicular bisector of the line PQ, we say that "the points P and Q are *symmetrically placed* with respect to the line l." Thus, the assertion of this problem is that $P(a, b)$ and $Q(b, a)$ are symmetrically placed with respect to the line $y = x$.

When two curves C_1 and C_2 are so situated that the points of the one and those of the other can be paired off so as to obtain pairs that are symmetrically placed with respect to a line l, we say that "the curves C_1 and C_2 are symmetrically placed with respect to the line l." A single curve C whose points can be paired off into pairs symmetrically placed with respect to a line l is said to be "symmetric with respect to l." For example, the curve represented by $y = x^2$ is symmetric with respect to the Y-axis, the locus of $y^2 = x$ with respect to the X-axis, that of $y + x = 5$ with respect to the bisector of YOX, that of $x^2 + y^2 = 9$ with respect to each of these three lines.

3. Prove that if $y = f(x)$ and $y = F(x)$ is a pair of inverse functions, then the graphs of these functions are symmetrically placed with respect to the line $y = x$.

4. Prove that the locus of the equation $x^2 + 2xy + y^2 = 4$ is a curve which is symmetric with respect to the line $y = x$.

5. Show the same thing for the curves represented by the following equations: (a) $x^3 + y^3 = 0$; (b) $3x^2 - 5xy + 3y^2 = 10$.

6. Prove that if $f(x, y) = 0$ represents a functional relation between x and y which is symmetric in x and y (compare p. 372), then the locus of this equation is symmetric with respect to the line $y = x$.

7. Determine the limit approached by

$$\text{(a)}\quad \frac{4}{2 + x} \text{ as } x \to -1$$

$$\text{(b)}\quad \frac{x + 3}{x - 1} \text{ as } x \to 3.$$

8. Compare the results obtained in *138*, 10 and 11 with the contents of Theorem LXXVIII.

9. Use this theorem to obtain the derivatives of the functions $y = x^{\frac{1}{n}}$, where n is a natural number.

10. Prove that $D_x e^{cx} = ce^{cx}$ (compare p. 374).

11. Prove the corollary to Theorem LXXX (compare p. 375).

12. Determine indefinite integrals of the following functions:

$$\text{(a) } \sin x; \quad \text{(b) } \cos 2x; \quad \text{(c) } e^{-x}; \quad \text{(d) } 2 \sin \frac{x}{3}.$$

13. Determine the area enclosed by the curve $y = \sin x$ from $x = 0$ to $x = \pi$, and the X-axis.

14. Show that the area enclosed by the curve $y = \frac{1}{x}$, the X-axis and

the ordinates $x = 1$ and $x = 3$ is equal to the area enclosed by the same curve, the X-axis and the ordinates $x = 5$ and $x = 15$.

15. Derive a general theorem concerning the curve $y = \frac{1}{x}$, of which 14 is a special case.

16. Prove that if the locus of the equation $y = f(x)$ is a simple plane curve, which lies in the 1st quadrant and whose slope is constantly positive, then the area enclosed by this curve, the Y-axis and the lines $y = b_1$ and $y = b_2$ $(b_2 > b_1)$ is equal to $a_2b_2 - a_1b_1 - \int_{a_1}^{a_2} f(x)\,dx$, when $b_1 = f(a_1)$ and $b_2 = f(a_2)$.

17. Determine the area bounded by the curve $y = e^x$, the Y-axis and the lines $y = e$ and $y = e^3$.

18. Use 3 to determine from 17 the area enclosed by the curve $y = \log_e x$, the X-axis and the lines $x = e$ and $x = e^3$.

19. Determine the derivative of $\tan x$ and of $\tan cx$, where c is a constant.

20. Show that if $y = f(x)$ and $x = F(y)$ is a pair of inverse functions, then the inverse of the function $y = f(cx)$ is $x = \frac{F(y)}{c}$.

21. Use the result of 20, in conjunction with Theorem LXXVIII, to prove that, if $D_xf(x) = f'(x)$, then $D_xf(cx) = cf'(cx)$.

22. Deduce the corollaries to Theorems LXXIX and LXXX from 21.

23. Set up a general formula for the kth derivative of $\sin x$ and of $\cos x$, in terms of the residue of k, mod. 4.

24. Interpret geometrically the fact that $D_x \sin x = \cos x$ and $D_x \cos x = -\sin x$, using the values of these functions given on page 100.

25. Discuss the motion of a point along the X-axis whose acceleration is proportional to its velocity, under the following initial conditions:

(a) When $t = 0$, the velocity is 0 and the point is at a distance a from O;

(b) When $t = 0$, the velocity is v_0 ft. per sec., and the point is at O;

(c) When $t = 0$, the point is at distance a from O and its velocity is v_0 ft. per sec.

26. Solve the following differential equations, the initial conditions being as indicated:

(a) $D_xy + 2y = 0$; $y = 5$ when $x = 0$;
(b) $D_{xx}y - 4y = 0$; $y = 5$ and $y' = 2$, when $x = 0$;
(c) $D_{xx}y + 4y = 0$; $y = 3$ and $y' = 4$, when $x = 0$.

27. A vertical steel spring, 10 in. in height, is fastened at the lower end, compressed to a height of 8 in. and then released. Suppose the mass of

the spring to be a unit, and the factor p^2 which enters in the elastic force to be equal to 4. Determine the velocity with which the free end of the spring will pass through its natural position, upward and downward.

28. Determine the equation of a curve through the point (1, 1) such that the triangle formed by any point P on the curve, the origin of coördinates and the point Q in which the tangent to the curve at P meets the X-axis, is an isosceles triangle with vertex at P.

29. Determine the equation of a curve through the point (0, 1) such that the distance from the point Q, in which the X-axis meets the tangent to the curve at an arbitrary point P, to the projection R of P on the X-axis is constantly equal to a constant a.

30. Determine a curve through the point (3, 4), such that the tangent to the curve at any point P is perpendicular to the line joining P to the origin of coördinates.

161. An important domain. We have now become somewhat acquainted with the character of differential equations, and we should therefore be able to absorb a few general ideas concerning them. It will have been noticed that these equations may differ as to the order of the derivatives which occurs in them. We distinguish correspondingly differential equations of the first, second or higher order according to the following definition (compare p. 370):

Definition LV. The *order* of a differential equation is the maximum of the orders of the derivatives of the unknown function which occur in it.

Thus, equations (14.1), (14.3) and (14.5) are of the first order; equations (14.2), (14.7) and (14.10) are of the second order.

As in algebraic equations, the degree of the power to which the unknown is raised is an important distinguishing mark in a differential equation. Since such equations involve not only an unknown function but also one or more of its derivatives, and since each of these can be raised to its own power, the definition of the degree of a differential equation is somewhat complicated.

We shall therefore introduce only a special case.

Definition LVI. A *linear* differential equation is a differential equation in which the unknown function and its derivatives enter only in terms of the first degree in all of them jointly.

It will be seen that all the differential equations which we have met so far are linear differential equations. Since nothing is said in the definition concerning the manner in which the independent

variable enters the equation, it follows that, for example, the equation

$$xD_{xx}y - \frac{2 + x}{x^2 - 1}D_xy + y \cdot \sin x = x^3$$

is also a linear equation, but that $yD_xy + x = 0$ and $D_xy - y^2 = 2x$ are not linear equations.

Newton's second law of motion which dominates such a large part of classical mechanics is responsible for the fact that differential equations of the second order play a particularly significant rôle in the mathematical study of natural science. For it brings the acceleration, and hence the second derivative with respect to time, into the analytical formulation of many problems. We shall therefore devote the remainder of this chapter to an important class of such equations, viz. to the linear differential equations of the second order.

162. The main highways. In accordance with Definitions LV and LVI the general form of linear differential equations of the second order is

$$(14.13)\qquad p_0(x) \cdot D_{xx}y + p_1(x) \cdot D_xy + p_2(x) \cdot y = q(x),$$

in which $p_0(x)$, $p_1(x)$, $p_2(x)$ and $q(x)$ are arbitrary functions of x. This general equation and several of its more or less specialized forms constitute the subject matter for a vast amount of literature.[1] We shall be occupied chiefly with a small part of this domain, viz. that in which the coefficients $p_0(x)$, $p_1(x)$ and $p_2(x)$ are constants. Thus we come to the "*linear differential equation of the second order with constant coefficients*":

$$(14.14)\qquad a_0 \cdot D_{xx}y + a_1 \cdot D_xy + a_2 \cdot y = q(x);$$

and in this set of equations, the subset for which $q(x) = 0$, i.e. the "linear *homogeneous* differential equation of the second order with constant coefficients,"

$$(14.14a)\qquad a_0 \cdot D_{xx}y + a_1 \cdot D_xy + a_2 \cdot y = 0,$$

plays an important part.

Many of the equations which we have encountered so far, in particular the important equation (14.7), are special instances of this type. It is from them that we obtain a suggestion for the

[1] The interested reader may take a look at the volume on *Modern Analysis*, by Whittaker and Watson, to obtain some idea of the character and extent of this literature.

solution of the equation (14.14a). But we shall begin with some general theorems which apply to the equation (14.13) and to the corresponding "homogeneous" equation.

$$(14.15) \qquad p_0(x) \cdot D_{xx}y + p_1(x) \cdot D_x y + p_2(x) \cdot y = 0,$$

under the special hypothesis that $p_0(x) \neq 0$ for all values of x under consideration.

Theorem LXXXI. If $y = u_1(x)$ satisfies equation (14.15), then $y = cu_1(x)$ also satisfies this equation, for every constant c.

The proof of this theorem is an immediate consequence of *138*, 24; the reader can carry out the details.

Theorem LXXXII. If $y = u_1(x)$ and $y = u_2(x)$ are solutions of the equation (14.15) then $y = u_1(x) + u_2(x)$ is also a solution of this equation.

This theorem is also very simple to prove; it follows from *138*, 22.

If we combine the contents of these two theorems, we obtain a result which contains them as special cases, viz.

Theorem LXXXIIa. If $y = u_1(x)$ and $y = u_2(x)$ are solutions of the equation (14.15) and c_1, c_2 are arbitrary constants, then $y = c_1u_1(x) + c_2u_2(x)$ is also a solution of this equation.

If in this theorem we put $c_1 = c_2 = 1$, we obtain Theorem LXXXII; from $c_1 = c$, $c_2 = 0$, we derive Theorem LXXXI.

The reader will wonder (at least he should wonder!) why we stop with *two* functions $u_1(x)$ and $u_2(x)$ in Theorem LXXXIIa; it would be equally true under the proper hypotheses that

$$y = c_1u_1(x) + c_2u_2(x) + c_3u_3(x)$$

or

$$y = c_1u_1(x) + c_2u_2(x) + c_3u_3(x) + c_4u_4(x)$$

are solutions. This question is answered in the following lemma and subsequent theorems.

Lemma. If $u_1(x)u'_2(x) - u'_1(x)u_2(x) = 0$ for all values of x in an interval (ab), then there exist two constants, c_1 and c_2, not both zero, such that $c_1u_1(x) + c_2u_2(x) = 0$ for all values of x in the interval.

Proof. (a) If $u_1(x) \equiv 0$,[1] the identity $c_1u_1(x) + c_2u_2(x) \equiv 0$ evidently holds for arbitrary choice of c_1 and for $c_2 = 0$.

[1] To designate that an equality, involving a variable x, holds for all values of x in an interval (ab), we use the notation $\equiv$, which is read "identically equal"; mention of the interval is omitted, when it is clear from the context which interval is meant. The negation of an identity is designated by the symbol $\not\equiv$. It is perhaps superfluous to add the remark that this use of the symbol $\equiv$ has no connection whatever with its use in congruences; see Chaps. VIII and IX.

(b) If $u_1(x) \neq 0$ for all values of x in the interval, it follows from the hypothesis that $\dfrac{u_1(x)u'_2(x) - u'_1(x)u_2(x)}{[u_1(x)]^2} \equiv 0$ and hence by means of *138*, 27 and Theorem LXXIII that $\dfrac{u_2(x)}{u_1(x)} \equiv c.$ But this means that $cu_1(x) - u_2(x) \equiv 0$, so that the required identity holds for $c_1 = c$ and $c_2 = -1$, the last of which is certainly not zero.

(c) If $u_1(x)$ vanishes for some, but not for all values of x in the interval, the proof involves more extended considerations; we have to take recourse to a reference to the books on analysis. It is easy to see that in this case $c_2 \neq 0$, so that the identity can then also be put in the form $u_2(x) \equiv cu_1(x)$.[1]

Theorem LXXXIII. If $u_1(x)$ and $u_2(x)$ are solutions of the differential equation $p_0(x)y' + p_1(x)y = 0$, in which $p_0(x) \neq 0$ in an interval (ab), then there exist two constants c_1 and c_2, not both zero, such that $c_1u_1(x) + c_2u_2(x) \equiv 0$; if $u_1(x) \not\equiv 0$, the identity can be put in the form $u_2(x) \equiv cu_1(x)$.

Proof. The hypothesis means that [2]

$$p_0u'_1 + p_1u_1 \equiv 0, \quad \text{and} \quad p_0u'_2 + p_1u_2 \equiv 0.$$

If we multiply the first of these identities by $-u_2$, the second by u_1 and add the results, we obtain

$$p_0(u_1u'_2 - u_2u'_1) \equiv 0;$$

since we suppose that $p_0 \neq 0$ throughout (ab), it follows that

$$u_1u'_2 - u_2u'_1 \equiv 0$$

in (ab). The lemma now supplies the desired conclusion.

Theorem LXXXIV. If $u_1(x)$, $u_2(x)$ and $u_3(x)$ are solutions of the equation (14.15) in which $p_0(x) \neq 0$ on (ab), no one of which is identically zero and such that there are no constants $\bar{c}_1$ and $\bar{c}_2$, not both zero, for which $\bar{c}_1u_1(x) + \bar{c}_2u_2(x) \equiv 0$, then there exist two constants, c_1 and c_2, not both zero, such that

$$u_3(x) \equiv c_1u_1(x) + c_2u_2(x).$$

[1] For the application of this lemma in the proofs of Theorems LXXXIII and LXXXIV, case (c) can be left out of consideration. In the general case, the lemma stated above is not correct. For a fuller treatment of these matters, the reader is referred to L. R. Ford, *Differential Equations*, pp. 66–69, 166–167.

[2] To simplify the notation, the variable x is omitted.

Proof. According to the hypothesis of the theorem, we have

$$p_0u''_1 + p_1u'_1 + p_2u_1 \equiv 0, \quad p_0u''_2 + p_1u'_2 + p_2u_2 \equiv 0$$

and

$$p_0u''_3 + p_1u'_3 + p_2u_3 \equiv 0.$$

If we multiply the first of these identities by $-u_2$, the second by u_1 and add the results, we find

$$(14.16a) \quad p_0(u_1u''_2 - u_2u''_1) + p_1(u_1u'_2 - u_2u'_1) \equiv 0.$$

In a similar manner, the first and third identities lead to

$$(14.16b) \quad p_0(u_3u''_1 - u_1u''_3) + p_1(u_3u'_1 - u_1u'_3) \equiv 0.$$

We introduce the abbreviations

$$(14.17) \quad v_1 \equiv u_1u'_2 - u_2u'_1 \quad \text{and} \quad v_2 \equiv u_3u'_1 - u_1u'_3.$$

By use of *138*, 26, we find then that

$$v'_1 \equiv u_1u''_2 + u'_1u'_2 - (u_2u''_1 + u'_2u'_1) \equiv u_1u''_2 - u_2u''_1,$$

and

$$v'_2 \equiv u_3u''_1 - u_1u''_3.$$

Hence the identities (14.16*a*) and (14.16*b*) take the form

$$p_0v'_1 + p_1v_1 \equiv 0 \quad \text{and} \quad p_0v'_2 + p_1v_2 \equiv 0 \text{ respectively.}$$

This shows that v_1 and v_2 satisfy the hypothesis of Theorem LXXXIII. Moreover $v_1 \not\equiv 0$ on (ab); for otherwise it would follow from (14.17) that $u_1u'_2 - u_2u'_1 \equiv 0$, in which case we would conclude from the lemma that there exist constants $\bar{c}_1$ and $\bar{c}_2$, not both zero, such that $\bar{c}_1u_1 + \bar{c}_2u_2 \equiv 0$, contrary to our hypothesis. Hence there exists a constant $-c_2$ such that

$$v_2 \equiv -c_2v_1, \quad \text{i.e.} \quad u_3u'_1 - u_1u'_3 \equiv -c_2(u_1u'_2 - u_2u'_1).$$

Rearrangement of terms reduces this identity to the form

$$u_1(u'_3 - c_2u'_2) - u'_1(u_3 - c_2u_2) \equiv 0.$$

To this relation, the lemma is again applicable, u_2 being replaced by the function $u_3 - c_2u_2$. Since moreover u_1 is by hypothesis not identically zero, there exists a constant c_1, such that

$$u_3 - c_2u_2 \equiv c_1u_1, \quad \text{i.e.} \quad u_3 \equiv c_1u_1 + c_2u_2.$$

Finally, c_1 and c_2 can not both be zero, since otherwise u_3 would vanish identically, contrary to the hypothesis. Thus our theorem is proved.

It shows that a solution of equation (14.15) of the form

$c_1u_1 + c_2u_2 + c_3u_3$, in which three different functions actually occur, could be reduced to a form in which only two are present; this would be true also of a solution in terms of four or more functions u_1, u_2, u_3, u_4, etc. It supplies the answer to the question which we put in the reader's mouth on page 386.

Corollary. The most general solution of the linear homogeneous differential equation of the second order (14.15) is given by $c_1u_1(x) + c_2u_2(x)$, where $u_1(x)$ and $u_2(x)$ are any two solutions of this equation for which no constants $\bar{c}_1$ and $\bar{c}_2$ exist, not both zero and such that $\bar{c}_1u_1(x) + \bar{c}_2u_2(x) \equiv 0$, and where c_1 and c_2 are arbitrary constants.

Every solution of the equation (14.15) can be put in this form. We can therefore say that this equation is completely solved as soon as two functions $u_1(x)$ and $u_2(x)$ have been found which satisfy the condition indicated in the corollary.

In this corollary, we find the justification of the answer given on page 378 to our question concerning the equation (14.7), which is a very special case of (14.15). The justification is incomplete on account of the gap in the proof of part (c) of the lemma.

163. A short bypath and a direct road. We proceed now with the general equation (14.13). Here we have two theorems of importance.

Theorem LXXXV. If $y = U_1(x)$ and $y = U_2(x)$ are two solutions of a linear differential equation of the second order, then $y = U_1(x) - U_2(x)$ is a solution of the corresponding homogeneous equation.

Proof. From

$$p_0U''_1 + p_1U'_1 + p_2U_1 \equiv q, \quad \text{and} \quad p_0U''_2 + p_1U'_2 + p_2U_2 \equiv q,$$

we obtain by subtraction the statement that

$$p_0(U''_1 - U''_2) + p_1(U'_1 - U'_2) + p_2(U_1 - U_2) \equiv 0,$$

i.e. that

$$p_0D_{xx}(U_1 - U_2) + p_1D_x(U_1 - U_2) + p_2(U_1 - U_2) \equiv 0.$$

This proves the theorem.

Theorem LXXXVI. If $y = U(x)$ is a solution of a linear differential equation of the second order and $y = u(x)$ is a solution of the corresponding homogeneous equation, then $y = u(x) + U(x)$ is also a solution of the linear differential equation. The proof can safely be left to the reader. From Theorems LXXXV and

LXXXVI in combination with our previous results, we obtain now the following conclusion:

Corollary. The most general solution of the equation (14.13) is given by $y = c_1u_1(x) + c_2u_2(x) + U(x)$, where $u_1(x)$ and $u_2(x)$ are two solutions of the corresponding homogeneous equation (14.15) which satisfy the conditions of the corollary to Theorem LXXXIV, $U(x)$ is any solution of the equation (14.13), and c_1 and c_2 are arbitrary constants.

The general theorems which have been developed carry us farther than is necessary for our immediate purpose. We have yielded in this instance to the temptation of wandering off the main road, because we had an opportunity to become acquainted, without much additional effort, with some of the basic facts in the domain of linear differential equations of the second order. We shall apply them to the special case of equation (14.14*a*), the linear homogeneous differential equation of the second order with constant coefficients, in order to obtain its general solution.

From the corollary of Theorem LXXXIV, we know that it suffices to find two solutions of this equation, which satisfy the condition stated there. The appearance of the exponential function in the solution of simple examples of this type, such as (14.7), leads us to inquire whether any function of the form e^{px} can be found to satisfy this more general equation. If $y = e^{px}$, then $D_xy = pe^{px}$ and $D_{xx}y = p^2e^{px}$. Therefore $y = e^{px}$ will satisfy equation (14.14*a*) if and only if p satisfies the equation

$$e^{px}(a_0p^2 + a_1p + a_2) = 0.$$

Since $e^{px} \neq 0$ for every finite value of x, this condition reduces to the algebraic equation

$$a_0p^2 + a_1p + a_2 = 0, \tag{14.18}$$

which we remember from many earlier occasions (compare e.g. p. 133). We know that this equation has in general two solutions, so that we have obtained all that is required for the application of Theorem LXXXIV and its corollary. This leads to the following result:

Theorem LXXXVII. If r_1 and r_2 are two distinct roots of the quadratic equation (14.18), then the general solution of the differential equation (14.14*a*) is given by $y = c_1e^{r_1x} + c_2e^{r_2x}$.

To justify this statement we still have to show that no con-

stants $\bar{c}_1$ and $\bar{c}_2$ exist, not both zero, for which $\bar{c}_1 e^{r_1 x} + \bar{c}_2 e^{r_2 x} \equiv 0$. If $\bar{c}_1$ were different from zero this identity would take the form $e^{(r_1 - r_2)x} \equiv -\frac{\bar{c}_2}{\bar{c}_1}$, which can certainly not hold if $r_1 \neq r_2$; if $\bar{c}_1 = 0$, the relation reduces to $\bar{c}_2 e^{r_2 x} \equiv 0$, which can hold only if $\bar{c}_2 = 0$ as well.

In a complete discussion of this theorem, we would now have to distinguish the cases in which the roots of the equation (14.18) are complex numbers from the case in which they are real numbers. Moreover the case in which the two roots coincide would require separate consideration. These matters we will have to leave for the more technical study of the subject.

Instead we shall consider a problem which approximates actual conditions more closely than any problem we have as yet dealt with. We return to the problem of the elastic body discussed in *159*, but we shall add the condition that the medium offers a resistance to the motion of the body, opposite in direction to the velocity of its free end, and proportional to it in magnitude. It should be clear from what has preceded, that in this case equation (14.10) should be replaced by the differential equation.

$$mD_{tt}x = -p^2x - q^2D_tx, \tag{14.19}$$

in which q^2 is a proportionality factor which depends on the physical character of the medium. The initial conditions remain as stated in (14.11). If we write b^2 in place of $\frac{q^2}{m}$ and proceed otherwise as on page 380, our problem reduces to

$$D_{tt}x + b^2D_tx + a^2x = 0,$$

which is indeed a linear homogeneous second order differential equation with constant coefficients. To apply Theorem LXXXVII, we have to solve for r the quadratic equation

$$r^2 + b^2r + a^2 = 0. \tag{14.20}$$

Its roots are found by use of (7.4) (see p. 133); we get

$$r_1 = \frac{-b^2 + \sqrt{b^4 - 4a^2}}{2} \quad \text{and} \quad r_2 = \frac{-b^2 - \sqrt{b^4 - 4a^2}}{2}.$$

From here on three cases would have to be considered, according as $b^4 - 4a^2$ is greater than, equal to, or less than zero. We shall

only consider the latter case. If we put $b^4 - 4a^2 = -n^2$ the roots of (14.20) take form

$$r_1 = -\frac{b^2}{2} + \frac{ni}{2}, \; r_2 = -\frac{b^2}{2} - \frac{ni}{2}.$$

We can now write down the general solution of the differential equation (14.19) in the form

$$x = k_1 e^{-\frac{tb^2}{2}} e^{\frac{tni}{2}} + k_2 e^{-\frac{tb^2}{2}} e^{-\frac{tni}{2}},$$

and hence, by means of (6.01) as on page 378, in the form

$$x = e^{-\frac{tb^2}{2}}\left[k_1\left(\cos\frac{nt}{2} + i\sin\frac{nt}{2}\right) + k_2\left(\cos\frac{nt}{2} - i\sin\frac{nt}{2}\right)\right];$$

i.e. $x = e^{-\frac{tb^2}{2}}\left(c_1\cos\frac{nt}{2} + c_2\sin\frac{nt}{2}\right),$

where $\qquad c_1 = k_1 + k_2$ and $c_2 = i(k_1 - k_2)$.

To find the velocity $D_t x$, we have to make use once more of *138, 26*, and of the corollaries to Theorems LXXIX and LXXX. The details are left for the reader's private enjoyment; he will find that

$$D_t x = e^{-\frac{tb^2}{2}}\left[\left(\frac{c_2 n}{2} - \frac{c_1 b^2}{2}\right)\cos\frac{nt}{2} - \left(\frac{c_1 n}{2} + \frac{c_2 b^2}{2}\right)\sin\frac{nt}{2}\right].$$

Finally we apply the initial conditions (14.11). We obtain the equations $d = c_1$ and $0 = \frac{c_2 n}{2} - \frac{c_1 b^2}{2}$, which lead to $c_1 = d$, $c_2 = \frac{db^2}{n}$; hence the solution of our problem is

$$x = e^{-\frac{tb^2}{2}}\left(d\cos\frac{nt}{2} + \frac{db^2}{n}\cdot\sin\frac{nt}{2}\right). \tag{14.21}$$

What does this solution mean?

The factor $e^{-\frac{tb^2}{2}}$ decreases as t increases; its rate of decrease diminishes as its magnitude gets smaller (compare p. 375), i.e. as t increases. The second factor represents an oscillatory movement; the character of this movement can be understood very readily if we make a slight transformation of this factor, based on the addition formulæ for the sine and cosine which we obtained in Chapter V (compare (5.16) on p. 100). By using these formulæ we find that $x = \frac{2ad}{n} e^{\frac{-tb^2}{2}} \cos\left(\alpha + \frac{nt}{2}\right)$, in which α is determined by the con-

ditions $\cos \alpha = \frac{n}{2a}$ and $\sin \alpha = -\frac{b^2}{2a}$; the reader should have no difficulty in showing that such an angle α exists and that this formula is indeed equivalent to (14.21). It is now seen that the factor $\frac{2ad}{n} \cdot \cos\left(\alpha + \frac{nt}{2}\right)$ and hence the second factor in (14.21) represents an oscillation of width $\frac{4ad}{n}$ and of period $\frac{4\pi}{n}$. The combined effect of the two factors is an oscillation of decreasing width (*amplitude* is the technical term), but of constant period. This type of motion is called *damped vibration;* it is indeed what a simple experiment shows to be the motion in a resisting medium of the free end of an elastic body, such as e.g. a steel spring, which has been extended or compressed. Thus we have finally reached a point at which the theoretical conclusions can be matched with occurrences in nature. We have only had a glimpse of this fitting of the theory into reality; perhaps it has been sufficient to give an idea of what lies beyond.

164. The physical universe. Let us now consider in retrospect the course through which this brief study of differential equations has led us. By means of Newton's second law of motion we were able to express in terms of differential equations of the second order the conditions that are imposed upon the motion of certain points in bodies acted upon by certain forces. These equations together with initial conditions determined the position and velocity of those points at all other times, past as well as future.

The problems which we have considered have been simple because the few forces which were involved acted along the same straight line. Their direction was constant and their magnitude depended on only one independent variable. In more complicated questions it is not quite so simple to represent the state of affairs by means of differential equations. Furthermore it is not always so simple a matter to obtain a solution of the differential equation. A good deal of technique has to be acquired by those who wish to be able to set up differential equations to represent more complicated situations and to draw intelligible conclusions from these differential equations. But the fundamental point of view is not essentially different from that which controls the procedure which has been sketched.

In the heyday of classical mechanics, it was thought that Newton's laws held the clue to the entire physical universe. To the famous French astronomer Pierre Simon de Laplace (1749–1827), who contributed a great deal to the application of these laws, and who taught us many things about the methods for extracting their inner secrets from differential equations, is attributed the boast that an intelligence which could know the position and velocity of all masses in the universe at some one instant and the forces acting on them, could then foretell the entire future development of this universe and reconstruct its past history.[1] The reader will see in this statement not so much a boast as a colorful way of asserting the universal validity of the laws of motion of Newton (or perhaps of some modification of these laws) and of the fact that a solution of a second order differential equation is determined when its value and its first derivative are known for one value of the independent variable.

As more and more phenomena are being brought under the sway of laws which can be formulated analytically in terms of differential equations, the significance of this subject becomes ever greater. It affords in this way a means for unlocking the secrets of the future. This presupposes that laws can be formulated for the control of phenomena, which remain unchanged for a period of time. The determination of such laws is usually considered to be the task of the scientist, physicist, chemist, biologist, economist(?); when they have been laid down in the form of differential equations, the task of drawing conclusions from them is put before the mathematician. For him it should be a matter of indifference what the source of these differential equations is: one and the same equation may arise from widely diverse domains. The mathematical results are then turned over to the scientist for interpretation and testing in his special field.[2]

[1] See Laplace, *Collected Works*, Vol. 7, p. VI; compare also Moritz, *Memorabilia Mathematica*, p. 328.

[2] For a clear discussion of the significance of the "laws" of science and their relation to mathematics the reader is referred to an excellent book by E. W. Hobson, *The Domain of Natural Science*.

CHAPTER XV

A REGION OF GREAT FAME

Beauty is truth, truth beauty, — that is all
Ye know on earth, and all ye need to know. — Keats.

165. Art in mathematics. Before very long our journey must come to an end. But we can not go home until we have had at least a glimpse of the region famed, wherever the life of the mind is held in esteem, for its beauty and seductive powers. In the field of projective geometry we possess a conquest of the human mind which has some of the qualities of imaginative art. While guided in many respects by concrete reality, it transcends, as poetry does, the bounds set by sensual perception and creates a world of its own. One would not have an adequate conception of mathematics if some experience with projective geometry had not gone into the making of it. From the field of differential equations which represents the domain of most intimate contact with the natural sciences we turn therefore to projective geometry upon which these other disciplines had until quite recent times not drawn for their own use. It is a significant sign that on the advanced front of the general theory of relativity, where generalized spaces are under consideration, the points of view and methods of projective geometry have recently become of considerable importance.

166.[1] More than the eye can see. From physiological optics we learn that the images our brain constructs by means of the eye are in many respects quite different from the "reality" they enable us to deal with. Rays of light emanating from various points of visible objects pass through the lenses of the eyes and are then cut by the retinas. It is from these retinal projections that optical images are formed. How the central nervous system deals with the retinal projections in order to form the visual impressions from

[1] Throughout this chapter we shall assume the postulates of Euclidean geometry (compare Chaps. X and XI). The reader will probably recognize most of the consequences of these postulates of which we have to make use; a textbook on plane geometry will be of assistance.

them, this is beyond our present range of interest; but we are concerned with the relation of the retinal projections of objects to their "objective reality." Let us begin with a very simple example. Suppose that on a line a we have points A, A_1, A_2, A_3 etc., *equally spaced*, and that we connect these points with a fixed point P (see Fig. 98). We can think of P as representing an eye, and of the lines AP, A_1P etc. as rays emanating from the points on the line a. Such "half-lines" through a point are frequently called *rays* by the analogy just suggested; their totality is called a *pencil of rays* (compare Definition LVIIb, p. 403). If we cut this pencil with a

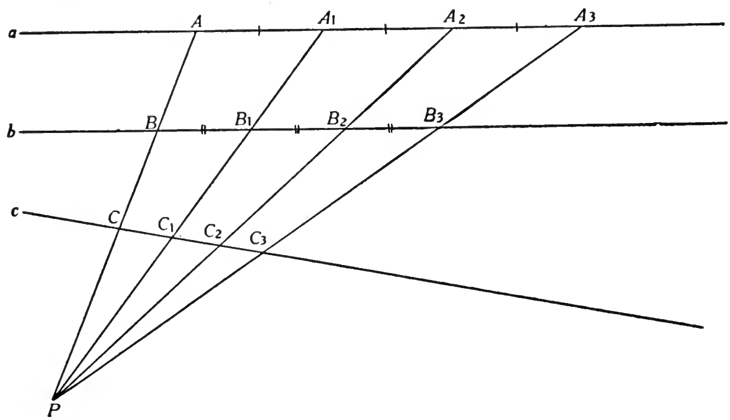


FIG. 98

line b, parallel to a, we obtain points B, B_1, B_2 etc., which in a certain sense can be thought of as images of the points A, A_1, A_2 etc., projected from P upon the line b; we can equally well think of the points on a as obtained by projecting the points on b from the fixed point P. How much information as to the points on a can we obtain from their projections on b; and vice-versa what can we conclude concerning the points on b from knowledge of those on a? In the *first* place, the distance between two points on the one line is *different* from the distance between the corresponding points on the other — this is a negative result. But *secondly*, from the equal spacing of the points on a, we can conclude that there is equal spacing of the points on b; and vice-versa — this is a positive result.[1]

[1] The reader will be able to prove this fact by using theorems in plane geometry with which he has been familiar since school days (see footnote on p. 395).

Let us now introduce a third line c, not parallel to a, and project the points of a and of b on it. The negative result of the previous case can also be stated in this case; but the positive result is no longer valid, for, if the points C, C_1, C_2 etc. were also equally spaced, c would be parallel to a. What *can* be affirmed concerning the points on c?

In a certain sense, this question is representative of the central problem of projective geometry — to determine those properties of geometric entities which are unaffected by projection. The geometric entities involved are not always quite as simple as the rows of points on the lines a, b and c; and "projection" includes other operations besides the one illustrated in Fig. 98. But the question remains the same. Let us see in how far we can answer it in the simple form which it takes here and how much we can learn from the answer.

It will perhaps occur to the reader when he meditates a little on the question and simply contemplates Fig. 98, that at least one thing appears to remain unchanged in passing from a to c, viz. the order in which the points follow each other; for, just as in passing from A to A_2, the point A_1 is encountered, so in moving from C to C_2, the point C_1 is met with. This is expressed simply by saying that when A_1 is *between* A and A_2, then the corresponding point C_1 is also *between* the points C and C_2 which correspond to A and A_2 respectively. And this seems to remain true, even if the points of a are projected on a line whose position with respect to a is like that of d in Fig. 99; although, in this case the direction in which the points succeed each other appears to be changed.

We might be tempted to conclude that, when the points of a line a are projected onto another line b, the betweenness relation for any three points on a is the same as that for the corresponding points on b. This would mean, in language which should be familiar to a reader of Chapter IV, that the set of points of a line a together with the relation of betweenness for sets of three among them is *isomorphic* with the set of points obtained by projecting the points of a on any other line, and the relation of betweenness for sets of three among *them* (compare pp. 64–65). Would this be true?

We remember that isomorphism requires in the first place equivalence, so that if the statement be true, projection from P

must determine a 1–1 correspondence between the points of a and those of any other line in the plane. But what are we to say of the point V_d on d (see Fig. 99), obtained as the intersection of d with a line through P parallel to a; what point on a corresponds to *it?* And, what point on d corresponds to the point V_a on a, for which V_aP is parallel to d? There is only one way out of this apparent difficulty — it would not have arisen at all if we had started in a strictly logical fashion by putting

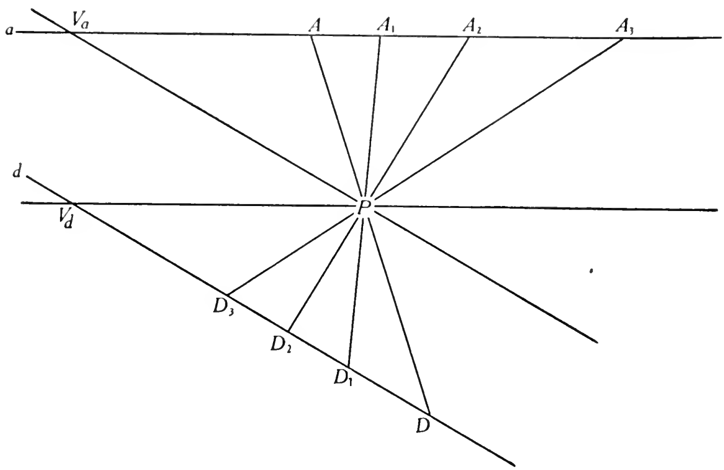


FIG. 99

down explicitly the primitive concepts and the postulates of projective geometry (compare *110–112*). We shall introduce piecemeal those for which the need arises. Here then are the first of them:

(15.1) Point and line are primitive concepts of projective geometry.

(15.2) Any two lines determine one and only one point.[1] This removes the difficulty, at least the logical difficulty, since on the basis of (15.2) the line through P and parallel to a also meets a in a single point. But it introduces a new difficulty, viz. that of

[1] It is to be understood that we are concerned only with the plane. Compare also the set of postulates on pp. 248–249.

the concrete representation of the elements of projective geometry, in particular of the points determined by pairs of parallel lines. We must equip each line with *a single exceptional point*, viz. *the point* in which it is met by all lines parallel to itself. This point is frequently called the *ideal point* of the line.

The traditional distinction between parallel and non-parallel lines is thus wiped out by the postulate of projective geometry given in (15.2); in other words, projective geometry recognizes no distinction between ideal points and other points. The terminology is a vestige of the historic fact that projective geometry grew up during the 19th century in an environment of ordinary geometry, usually called *metric* geometry when the distinction from projective geometry is to be emphasized (compare p. 421). The subject was not established on a strictly logical basis until recent times. Adhering to the historic terms we can answer the question that was raised a moment ago by saying: To the point V_d on d corresponds the ideal point on a; to the point V_a on a corresponds the ideal point on d.

Along with postulate (15.2) attention must be called to the following postulate:

(15.2*a*) Any two points determine one and only one line,

and to the fact that (15.2) and (15.2*a*) are obtainable one from the other by interchanging the words *point* and *line*.

As to the isomorphism by which we were tempted above, we can now say that it does not hold. For, the betweenness relation for three points on a (see Fig. 99) is the same as that for the corresponding points on d only if all three points lie on the same side of V_a.

167. New experiences and novel sounds. We can now deal a little more fully with the general question of projective geometry, as applied to our simple problem. We have already seen that the distance between two points on a line is not invariant under projection. We will see presently that for three points on a line a the ratios of their mutual distances are invariant if the points are projected on a line parallel to a, such as b in Fig. 98, but not if they are projected on an arbitrary line, like c. But if four points are given on a line it is possible to calculate a number which does remain invariant under projection on an arbitrary line. To establish these facts we have to prove first an interesting theorem,

known as the *Law of Sines* to every one who has studied trigonometry. A few preliminary facts are necessary.

1. From the definitions of $\sin\theta$ and $\cos\theta$ (see p. 98) and the interpretation of positive and negative angles as respectively counter-clockwise and clockwise rotations (compare p. 97), we conclude at once that

$$(15.3) \quad \cos(-\theta) \equiv \cos\theta, \quad \text{and} \quad \sin(-\theta) \equiv -\sin\theta.^{1}$$

2. If we substitue π for θ_1 and $-\theta$ for θ_2 in the addition formulas (5.16) of page 100, we obtain

$$\cos(\pi - \theta) \equiv \cos\pi\cos(-\theta) - \sin\pi\sin(-\theta), \text{ and}$$

$$\sin(\pi - \theta) \equiv \sin\pi\cos(-\theta) + \cos\pi\sin(-\theta).$$

Hence, remembering that π is the radian measure of an angle of 180°, using the results of *54* (see p. 100) and also formulas (15.3), we find that

$$(15.4) \quad \cos(\pi - \theta) \equiv -\cos\theta, \quad \text{and} \quad \sin(\pi - \theta) \equiv \sin\theta,$$

i.e. *the sine of an angle equals the sine of its supplement, the cosine of an angle is equal to the negative of the cosine of its supplement.*

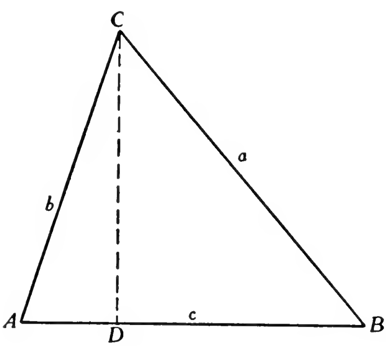


FIG. 100

We are now ready to prove

Theorem LXXXVIII. (Law of Sines.) The ratio of the sines of any two angles of a triangle is equal to the ratio of the sides opposite them.

Proof. (a) If the two angles are acute, as angles A and B in Fig. 100, the proof of this theorem follows immediately from a consideration of the two right-angled triangles ADC and BDC. From the former we obtain $\sin A = \dfrac{DC}{AC}$; and from the latter $\sin B = \dfrac{DC}{BC}$. Hence, denoting the length of a side of the triangle by the small letter corresponding to the capital letter at the opposite vertex, we have:

[1] Compare footnote 1 on p. 386.

$$DC = b \sin A, \quad \text{and} \quad DC = a \sin B,$$

and therefore $b \sin A = a \sin B$, or $\dfrac{\sin A}{\sin B} = \dfrac{a}{b}$.

(b) If one of the angles is obtuse, as in Fig. 101, we consider the right triangles ABD and CBD. We find then: $\sin A = \dfrac{DB}{c}$, and $\sin(\pi - C) = \dfrac{DB}{a}$; hence $DB = c \sin A$, and $DB = a \sin(\pi - C)$. From this, in connection with (15.4), we derive the fact that $\dfrac{\sin A}{\sin C} = \dfrac{a}{c}$.

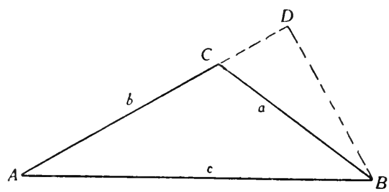


FIG. 101

We have therefore established the general validity of the *law of sines*, viz.: between the angles and sides of any triangle, the continued proportion

$$\sin A : \sin B : \sin C = a : b : c \tag{15.5}$$

holds true.

We return then to the consideration of points on a line. If we apply the law of sines to $\triangle A_1PA_2$ (see Fig. 102), we find that $\dfrac{A_1A_2}{A_2P} = \dfrac{\sin A_1PA_2}{\sin PA_1A_2}$; and from $\triangle A_2PA_3$ we obtain that $\dfrac{A_2A_3}{A_2P}$

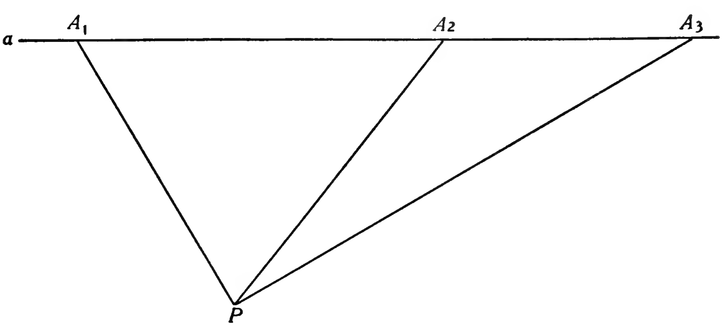


FIG. 102

$= \dfrac{\sin A_2PA_3}{\sin PA_3A_2}$. Dividing the sides of the first of these equations by the corresponding sides of the second, it follows that

$$\frac{A_1A_2}{A_2A_3} = \frac{\sin A_1PA_2}{\sin A_2PA_3} \times \frac{\sin PA_3A_2}{\sin PA_1A_2}.$$

But from $\triangle PA_1A_3$ comes the relation

$$\frac{\sin PA_3A_2}{\sin PA_1A_2} = \frac{PA_1}{PA_3},$$

so that the preceding equation takes the form

$$\frac{A_1A_2}{A_2A_3} = \frac{\sin A_1PA_2}{\sin A_2PA_3} \times \frac{PA_1}{PA_3}. \tag{15.6}$$

If the points of a are now projected from P upon an arbitrary line b, we shall find in the same way that

$$\frac{B_1B_2}{B_2B_3} = \frac{\sin B_1PB_2}{\sin B_2PB_3} \times \frac{PB_1}{PB_3}. \tag{15.6a}$$

From (15.6) and (15.6a), it follows that $\dfrac{A_1A_2}{A_2A_3}$ is equal to $\dfrac{B_1B_2}{B_2B_3}$ *if and only if* b is parallel to a, as was asserted in the first paragraph of this section (compare p. 399). Furthermore we observe that the last ratio occurring in (15.6) is entirely independent of the position of the point A_2. If therefore we take a fourth point A_4 on a (see Fig. 103), we shall find:

$$\frac{A_1A_4}{A_4A_3} = \frac{\sin A_1PA_4}{\sin A_4PA_3} \times \frac{PA_1}{PA_3}.$$

Division of the two sides of this equation into the corresponding sides of (15.6) leads therefore to the following important conclusion:

$$\frac{A_1A_2}{A_2A_3} \div \frac{A_1A_4}{A_4A_3} = \frac{\sin A_1PA_2}{\sin A_2PA_3} \div \frac{\sin A_1PA_4}{\sin A_4PA_3}. \tag{15.7}$$

Several conclusions can be derived from this result. We observe that the right side depends only on the relative position of the four lines through P, but is independent of the position of the line a; and that the left side depends only on the relative position of the points on the line a, but is independent of the position of the point P. From the first of these observations it follows that if

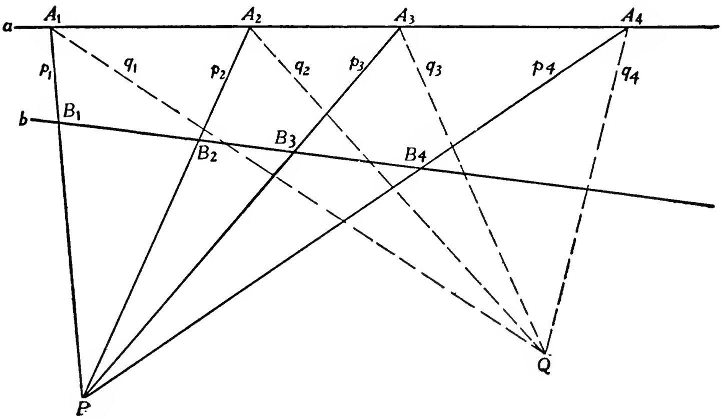


FIG. 103

the points of a are projected from P on an arbitrary line b, and an expression analogous to the left side of (15.7) is calculated for the points B_1, B_2, B_3, B_4, this expression will turn out to be equal to the right side of (15.7). From the second observation we conclude that if the points A_1, A_2, A_3, A_4 are connected with a new point Q, and an expression analogous to that on the right side of (15.7) be calculated for the point Q, there will result the left side of (15.7). In other words, the expression $\dfrac{A_1A_2}{A_2A_3} \div \dfrac{A_1A_4}{A_4A_3}$ is invariant under projection from a point; and the expression $\dfrac{\sin \widehat{p_1p_2}}{\sin \widehat{p_2p_3}} \div \dfrac{\sin \widehat{p_1p_4}}{\sin \widehat{p_4p_3}}$, in which $\widehat{p_1p_2}$ designates the angle between lines p_1 and p_2, has the same value if calculated for two sets of concurrent lines, in which the points of intersection of corresponding lines of the two sets lie on a straight line, as e.g. the lines PA_i and QA_i, $i = 1, 2, 3, 4$ in Fig. 103. But more valuable and interesting material is within our reach; its consideration is greatly facilitated by the use of a few technical terms and notations, some of which we have met before (see p. 396) — hence, some definitions.

Definition LVIIa. A set of points on a line is called a *pencil of points;* the line is called the *base* of the *pencil.*

Definition LVIIb. A set of lines on (through) a point is called a *pencil of lines;* the point is called the *vertex* of the pencil.

Definition LVIIc. The word (*plane*) *pencil* is used to designate either a pencil of points or a pencil of lines.

Definition LVIIIa. The number $\frac{A_1A_2}{A_2A_3} \div \frac{A_1A_4}{A_4A_3}$ calculated for four points of a pencil of points is called the *double ratio* (or *cross ratio*, or *anharmonic ratio*) of the four points and is denoted by $(A_1A_2A_3A_4)$.

Definition LVIIIb. The number $\frac{\sin \widehat{p_1p_2}}{\sin \widehat{p_2p_3}} \div \frac{\sin \widehat{p_1p_4}}{\sin \widehat{p_4p_3}}$ calculated for four lines of a pencil of lines is called the double ratio (cross ratio, anharmonic ratio) of the four lines and is denoted by the symbol $(p_1p_2p_3p_4)$.

Definition LIXa. If the lines joining corresponding points of two pencils of points meet on a point, the pencils are said to be *in perspective;* the notation, as applied to Fig. 103, is $A_1A_2A_3A_4 \overset{=}{\wedge} B_1B_2B_3B_4$.

Definition LIXb. If the points joining corresponding lines of two pencils of lines meet on a line, the pencils are said to be in perspective; the notation, as applied to Fig. 103, is $p_1p_2p_3p_4 \overset{=}{\wedge} q_1q_2q_3q_4$.

Definition LIXc. If the points of a pencil of points lie on the corresponding lines of a pencil of lines, the two pencils are said to be in perspective; the notation, as applied to Fig. 103, is $A_1A_2A_3A_4 \overset{=}{\wedge} q_1q_2q_3q_4$.

Four points are to be noticed about these definitions:

1. The two definitions *a* and *b* in each set are so related that either of them is obtainable from the other by interchanging the words point and line.

2. To obtain the relation just described in a strict sense, some words had to be used in an unusual manner: we have spoken, e.g., of two lines "meeting *on* a point," instead of "meeting *in* a point" which is the usual form, and of the "point joining two lines" whereas the customary language speaks of the "point of intersection of two lines." In our further discussion it may not always be convenient to strain the language in this way; the reader should have no difficulty in making the necessary transition from the conventional terminology to the forms that are necessary to obtain a strict point $\longleftrightarrow$ line correspondence.[1]

3. Besides making departures in language, we had to replace

[1] Compare also the remark made at the end of *166*.

"distance between points" by "sine of the angle between lines," as in Definitions LVIII*a* and LVIII*b*, to obtain this strict correspondence.

4. What we have called thus far "projecting points of a line, from a point, on another line" is seen to be covered by the phrase "passing from a pencil of points to a perspective pencil of points."

By the aid of these definitions we can state the outcome of our discussion as follows:

Theorem LXXXIX. The double ratios of two sets of four elements each are equal if these sets consist of corresponding elements of two perspective pencils.

This result must suffice for the present — it gives a partial answer to the question referred to on page 397 as representative of the central problem of projective geometry; incidentally we have acquired some other ideas — have we?

168. To make the novelties familiar.

1. Show that, in Fig. 98, either of the equalities $AA_1 = A_1A_2$ and $BB_1 = B_1B_2$ is a consequence of the other.

2. Prove that if a pencil of lines is cut by two parallel lines, a and b, then the segments cut off on b are proportional to the segments cut off on a. (Evidently 1 is a special case of this theorem.)

3. Three points, A, A_1 and A_2 are given on a line a, so that $AA_1 = A_1A_2$. They are connected by straight lines with a point P, and a point B is taken on AP, so that $AB = BP$. Construct a line b through B so that $BB_1 = \frac{B_1B_2}{2}$, when B_1 and B_2 are the points of intersection of b with PA_1 and PA_2 respectively.

4. Generalize the construction of problem 3 to the case in which the ratio $\frac{BB_1}{B_1B_2}$ is equal to an arbitrary rational number instead of being equal to $\frac{1}{2}$.

5. The point A is given on the line a. The pencil of points on a is to be projected, from the given point P, on a line b. Construct this line b in such a way that in the projection A will correspond to the ideal point on b.

6. Show that if A_1, A_2 and A_3 are three points arbitrarily placed on a directed line a, then $A_1A_2 + A_2A_3 + A_3A_1 = 0$.

7. Prove the following corollary to the theorem proved in 6. If P, Q, R are three points arbitrarily placed on a directed straight line a, then $PQ = PR + RQ$.

8. Prove that, if A_1 and A_2 are given points, there exists one and only

one point P on the directed line determined by them such that $\frac{A_1P}{PA_2} = 1$.

9. Prove that for no finite point, Q, on the directed line determined by A_1 and A_2 is the ratio $\frac{A_1Q}{QA_2}$ equal to -1.

10. Derive from the theorem proved in 6, that, as A_3 tends toward the ideal point on a, the ratio $\frac{A_1A_3}{A_3A_2}$ tends to -1; and hence, in connection with 9, that, if $\frac{A_1A_3}{A_3A_2} = -1$, the points A_1 and A_2 must be at finite distance and the point A_3 the ideal point on the line determined by them.

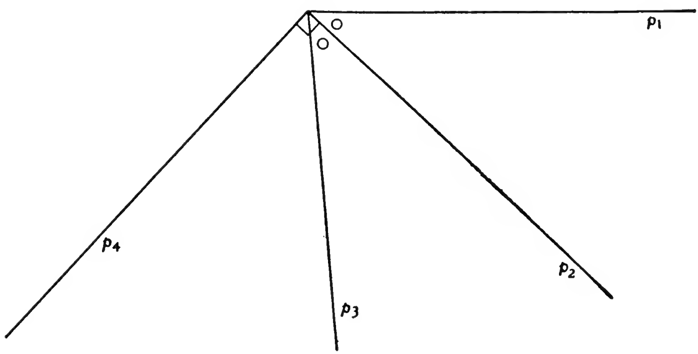


FIG. 104

11. Prove that if A_1, A_2, A_3 and A_4 are four points arbitrarily placed on a directed line a, then the double ratio $\frac{A_1A_2}{A_2A_3} \div \frac{A_1A_4}{A_4A_3}$ tends to a definite limit as the point A_4 tends toward the ideal point on a; determine this limit.

12. Prove that the double ratio of the four points (or lines) of a pencil is changed to its reciprocal value if the 1st point is replaced by the 2nd, the 2nd by the 3rd, the 3rd by the 4th, and the 4th by the 1st.

13. Prove that the double ratio of four points is not changed by reversing the order of the points.

14. Are there any other changes in the relative position of four points besides the one mentioned in 13, which do not affect the value of the double ratio?

15. Determine the double ratio of the four lines p_1, p_2, p_3 and p_4 in Fig. 104; p_2 is the bisector of the angle $\widehat{p_1p_3}$, and p_4 is perpendicular to p_2.

16. Show that a set of parallel lines is a special case of a pencil of lines.

17. Devise a method for determining the double ratio of a set of four parallel lines; apply it to determine the double ratio of four equally spaced parallel lines, taken in the order in which they are placed.

18. If the sides AB and AC of a triangle ABC are designated by p_1 and p_3 respectively, the bisector of the angle A by p_2 and the median through A by p_4, determine the double ratio of the pencil $p_1p_2p_3p_4$.

19. The lines p_1, p_2 and p_3 being the same as in 15, but p_4 the altitude through A, determine the cross-ratio of the pencil $p_1p_2p_3p_4$.

20. Determine the anharmonic ratio $(p_1p_2p_3p_4)$ when p_1 and p_2 are two sides of a triangle, p_3 the median and p_4 the altitude line through their common point.

169. At the seat of power. It follows from Theorem LXXXIX that if two pencils are so related that, although not in perspective themselves, each of them is in perspective with a third pencil, the double ratios of these two pencils will still be equal. More generally, if two pencils, let us denote them simply by the single letters π_1 and π_2, are so related that there exist a finite number of auxiliary pencils $\pi_3, \pi_4, \ldots \pi_k$, such that

$$\pi_1 \overset{=}{\wedge} \pi_3 \overset{=}{\wedge} \pi_4 \overset{=}{\wedge} \cdots \overset{=}{\wedge} \pi_k \overset{=}{\wedge} \pi_2,$$

then the double ratios of π_1 and π_2 will also be equal; we say in such a case, that π_1 and π_2 are connected by a "chain of perspectivities." An example is furnished by Fig. 105, in which $A_1A_2A_3A_4 \overset{P}{\underset{\wedge}{=}} B_1B_2B_3B_4 \overset{Q}{\underset{\wedge}{=}} C_1C_2C_3C_4$.[1] The following definition embodies this type of relation.

Definition LX. If for two pencils a finite number of auxiliary pencils exist which join the two given pencils by a chain of perspectivities, they are said to be *projectively related;* the relation is called a projectivity and is denoted by the symbol $\overline{\wedge}$.

For example, in Fig. 105,

$$A_1A_2A_3A_4 \;\overline{\wedge}\; C_1C_2C_3C_4 \quad \text{and} \quad p_1p_2p_3p_4 \;\overline{\wedge}\; q_1q_2q_3q_4.$$

As an immediate consequence of Theorem LXXXIX, we can state the following result:

Theorem XC. The double ratios of two sets of four elements

[1] It is frequently convenient, in stating that two pencils of points (lines) are in perspective, to indicate the point of intersection of the lines which join corresponding points (the line joining the points of intersection of corresponding lines). This is done as indicated in the text.

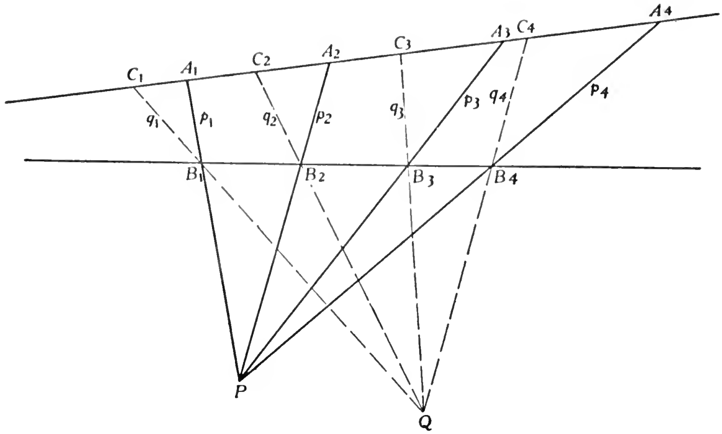


FIG. 105

each, consisting of corresponding elements of two projectively related pencils, are equal.

It is important to observe that, whereas the perspectivity of two pencils of points (lines) can occur only if they have different bases (vertices), two projective pencils of points (lines) may very well have the same base (vertex); Fig. 105 furnishes illustrations of this fact.

The definition of a projectivity admits so many possibilities that one is led to wonder whether it may not be true that any two arbitrary pencils are projectively related; e.g., whether, a correspondence between two pencils of points having been given arbitrarily, it is not always possible to determine other pencils of points which would connect the first two by a chain of perspectivities. This is not the case. In how far may a 1–1 correspondence between two pencils be chosen arbitrarily if they are to be projectively related? To answer this question, some preparation is needed. It is supplied by the following simple theorems:

Theorem XCI. If on two lines a and b (see Fig. 106) two pairs of points, A_1 and B_1, A_2 and B_2 are given, there always exists one and only one perspectivity in which these points are corresponding points.

Proof. The points A_1 and B_1 determine a single line, by (15.2*a*); let us denote it by A_1B_1. Similarly the points A_2 and B_2 determine

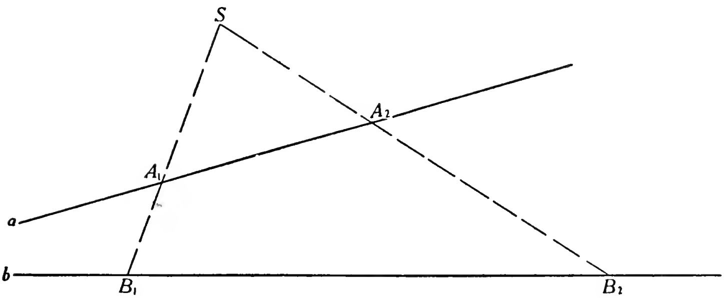


FIG. 106

a line, A_2B_2. The lines A_1B_1 and A_2B_2 determine a single point S, by (15.2). Projection of the pencil of points on a from S onto the line b establishes a perspectivity with the pairs of corresponding points A_1, B_1 and A_2, B_2. If there were another perspectivity with these same pairs of corresponding points, there would be another point common to the lines A_1B_1 and A_2B_2; this would be contrary to (15.2).

Theorem XCII. If two points, P and Q, and the real number r are given, there exists one and only one point R on the line PQ such that the ratio $\dfrac{PR}{RQ}$ of the directed segments PR and RQ is equal to r.

Proof. (See Fig. 107.) Through the point Q we draw a line segment QQ' arbitrary in direction and arbitrary in length; through P we draw a line $P'PP''$ parallel to QQ'. If the given real number r is positive, we determine on the former line a point P', on the opposite side of PQ from Q', such that, with the line-segment QQ' as a unit, the point P' corresponds to r. The point in which the lines $P'Q'$ and PQ meet is the required point R. If r is negative, we determine a point P'' on the same side of PQ as Q', which corresponds to $|r|$, the segment QQ' again being the unit; the common point of $P''Q'$ and PQ gives then the required point R_1. The former construction always gives a point R between P and Q, so that the segments PR and RQ are of the same sign and their ratio is positive. That $\dfrac{PR}{RQ} = r$ follows from the similarity of the triangles PRP' and QRQ'. In the second construction, the point R_1 will always lie outside the segment PQ (on the side of Q if $|r| > 1$,

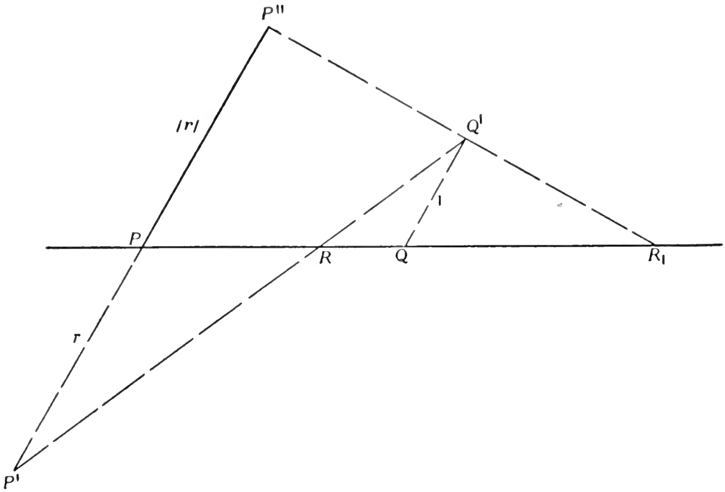


FIG. 107

on the side of P if $| r | < 1$), so that the segments PR and RQ are of opposite sign and this ratio is negative. The similarity of the triangles PR_1P'' and QR_1Q' proves that $\dfrac{PR_1}{R_1Q} = r$.

Hence we can always find *a* point R, which satisfies the requirements of the theorem. Are there more such points? Suppose that

$$\frac{PR}{RQ} = \frac{PS}{SQ}, \tag{15.8}$$

the points R and S being on the directed line determined by P and Q. This relation would, by means of a simple multiplication and by use of *168*, 7 be equivalent to $(PS + SR)SQ = PS(RS + SQ)$, thus to $SR \cdot SQ = PS \cdot RS$, and finally to $RS(PS - QS) = 0$. There are therefore two possible consequences of the assumption (15.8), viz. *either* $RS = 0$, which means that R and S coincide; or else, $PS = QS$, from which, in conjunction with (15.8) itself, we would conclude, since P and Q are distinct points, that $\dfrac{PR}{RQ} = \dfrac{PS}{SQ}$ $= -1$. But this would mean in view of *168*, 10, that both R and S are the ideal point on the line PQ, so that again they coincide.

Thus we have shown that there can be *only one* point R for which $\frac{PR}{RQ} = r$; this completes the proof of Theorem XCII.

Remark. It should be observed that the constructions in Fig. 107 are very similar to those carried out in *41* (pp. 72–75) and that the proof of Theorem XCII uses in an essential way the assumption of Chapter V (compare p. 72). This assumption as to the isomorphism of real numbers and points on a line should therefore be counted among the postulates of projective geometry as here developed.

Theorem XCIII. If three collinear points P, Q, R and a real number r are given, there exists one and only one point S such that the double ratio $\frac{PR}{RQ} \div \frac{PS}{SQ} = r$.

Proof. This theorem is an immediate consequence of the preceding one. For, from the fact that the double ratio is equal to r, we conclude that $\frac{PS}{SQ} = \frac{PR}{RQ} \div r$, so that S is always uniquely determined, when P, Q, R and r are given. When $r = 1$, and in that case only, S will coincide with R.

Corollary. The double ratio of four distinct points can never be equal to 1.

Furthermore, it follows from Theorem XCIII that if three collinear points P, Q, R and a real number r are given, then there always exists a single fourth point S, such that the double ratio of the four points, *taken in any other order* than that used in the theorem, is equal to r. For a complete proof of this statement, detailed answers to *168*, 12, 13 and 14 are useful. But even without these answers, the reader will see that, if the double ratio of P, Q, R and S is given for any one of the 24 possible orders of these points, the number $\frac{PR}{RQ} \div \frac{PS}{SQ}$ is thereby determined (even though he may not know how to determine it), so that Theorem XCIII becomes applicable.

We are now ready to answer the question put on page 408.

Theorem XCIV. If three pairs of elements of two pencils of points on different bases are given, one and only one projectivity between the pencils is thereby determined, in which these pairs consist of corresponding points.

Proof. Suppose that we consider the pencils of points on the two lines a and b, and that the pairs of points, A_1 and B_1, A_2 and B_2, A_3 and B_3 are given (see Fig. 108). We have then to show (1) that

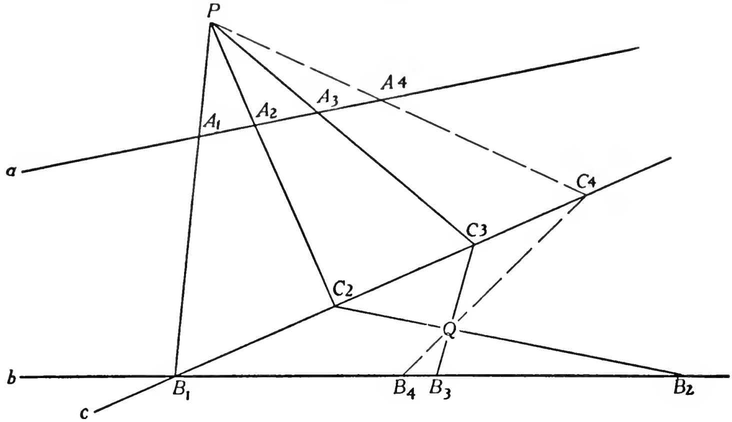


FIG. 108

there exists a chain of perspectivities in which the two given pencils are the terminal pencils, having the points of the given pairs as corresponding points; and (2) that, no matter what chain

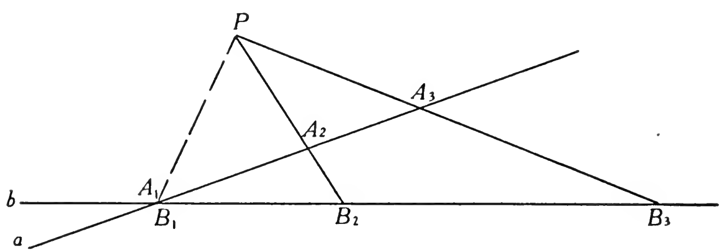


FIG. 109

of perspectivities may be used for this purpose to an arbitrary point of one of the given pencils, there corresponds a single point of the other. To prove the first of these assertions, we take an arbitrary point P on the line determined by B_1 and A_1,[1] and we draw an

[1] If A_1 and B_1 are coincident, as in Fig. 109, the two pairs of points A_2 and B_2, A_3 and B_3 determine a perspectivity, as proved in Theorem XCI; in this perspectivity: A_1B_1, A_2B_2 and A_3B_3 are pairs of corresponding points. We may therefore suppose that A_1 and B_1 do not coincide.

arbitrary line c through B_1. We project the points A_2, A_3 from P on c, so as to obtain the points C_2 and C_3. Finally we determine the point Q in which the lines C_2B_2 and C_3B_3 meet. We have then

$$A_1A_2A_3 \overset{P}{\overline{\overline{\wedge}}} B_1C_2C_3, \quad \text{and} \quad B_1C_2C_3 \overset{Q}{\overline{\overline{\wedge}}} B_1B_2B_3.$$

This represents the chain (a very short one) of perspectivities which was required. To the point A_4 on a corresponds the point B_4 on b obtained by projecting A_4 from P on c, and C_4 from Q on b; and conversely.

As to the second assertion: if A is an arbitrary point on a, to which correspond the points B and B' in two different projectivities, in which A_1 and B_1, A_2 and B_2, A_3 and B_3 are corresponding pairs, then we know from Theorem XC that the double ratios of $B_1B_2B_3B$ and of $B_1B_2B_3B'$ are both equal to $(A_1A_2A_3A)$ and hence equal to each other, so that, in virtue of Theorem XCIII, B and B' must coincide. This completes the proof of our theorem.

It is now an easy matter and a very tempting occupation to deduce from this theorem a great many consequences; the reader will have an opportunity to do this a little later on (see *171*). In particular, it can readily be shown that the theorem obtained by the interchange of the words point and line in Theorem XCIV is also valid (compare 2 on p. 404). Our immediate aim has been reached in so far as we have shown that, if more than three pairs of corresponding elements of two pencils of points are given, the pencils are not projectively related if the correspondence of elements is not of such character that the double ratio of any four elements of either pencil equals that of the corresponding elements of the other.

If, on the other hand, the correspondence between two pencils of points is such that the double ratio of any four elements of one pencil is equal to that of the corresponding elements of the other, then the two pencils are in projective relation. Let A_1, A_2, A_3 etc. and B_1, B_2, B_3 etc. be corresponding points of the two pencils. Then, as has just been proved, a *unique* projectivity is determined by the correspondence of A_1 to B_1, A_2 to B_2, and A_3 to B_3. In this projectivity there corresponds to any fourth point A a unique point B', such that $(A_1A_2A_3A) = (B_1B_2B_3B')$. But the point B, which is associated with A in the given correspondence between the pencils, is, by hypothesis, such that $(A_1A_2A_3A) = (B_1B_2B_3B)$, so that $(B_1B_2B_3B') = (B_1B_2B_3B)$. Hence, in virtue of Theorem

XCIII, B' coincides with B; i.e. the given correspondence is indeed a projectivity. We record this result in the following fundamental theorem.

Theorem XCV. A correspondence between two pencils of points which leaves the double ratio of any four elements invariant is a projectivity.

We turn now from these general considerations to have a look at sets of elements for which the double ratio has a special value. In this direction we have only obtained so far the negative result contained in the corollary to Theorem XCIII.

170. Proving the title to fame. We have seen in the proof of Theorem XCII and also in *168*, that to the value -1 of the ratio $\frac{A_1A_2}{A_2A_3}$ corresponds a somewhat exceptional position of the point A_2, viz. that of the ideal point on the line A_1A_3. It should not be surprising that the value -1 for the double ratio of four elements of a pencil is also of especial interest (compare *168*, 15). It indicates in the case of four points of a pencil of points that two of the points divide the segment determined by the other two internally and externally in ratios of the same numerical value; this is known, in elementary geometry, as harmonic division of a segment. Accordingly we have the following definition.

Definition LXI. If four elements of a pencil are so placed that their double ratio, in some order, is equal to -1, they are said to form a *harmonic pencil.*

If the elements are, in general, e_1, e_2, e_3 and e_4 and their double ratio is $\frac{e_1e_2}{e_2e_3} \div \frac{e_1e_4}{e_4e_3}$, then e_2 and e_4 are called *harmonic conjugates* with respect to e_1 and e_3, and e_1 and e_3 *harmonic conjugates* with respect to e_2 and e_4.

A simple example of a harmonic pencil of points is formed by two vertices of a triangle and the points in which the side joining them is met by the internal and external bisectors of the opposite angle. Either pair of points consists of harmonic conjugates with respect to the other pair. Of greater interest are the "configurations" known as the "complete quadrilateral" and the "complete quadrangle."

Definition LXII. A configuration is a collection of points and lines.

Definition LXIIIa. A complete quadrilateral is the configuration consisting of four distinct lines and the six distinct points determined by them. The four lines are called the *sides* of the quadrilateral, their six points of intersection are the *vertices*, the three lines, apart from the sides, determined by the vertices, are called the *diagonals*.

Definition LXIIIb. A complete quadrangle is the configuration consisting of four distinct points and the six distinct lines determined by them. The four points are called the *vertices* of the

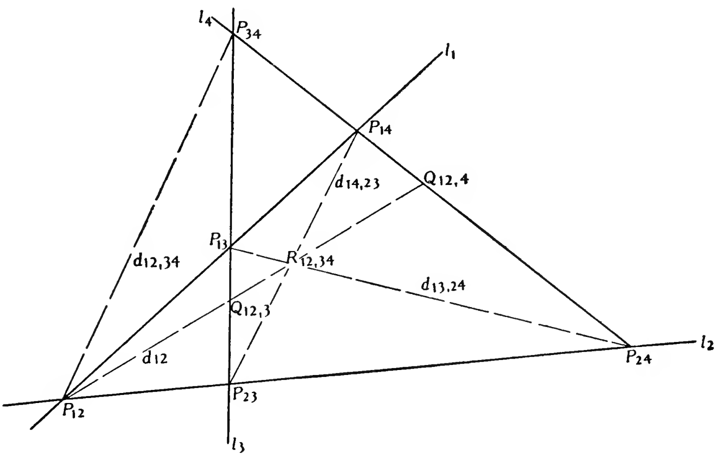


FIG. 110

quadrangle, their six lines of juncture the *sides*, and the points, apart from vertices, determined by the sides, are the *diagonal points*.[1]

We can prove now the following remarkable theorem:

Theorem XCVI. Two sides of a complete quadrilateral are harmonic conjugates with respect to the diagonal through the vertex determined by them and the line joining this vertex to the point of intersection of the other diagonals; these four lines form a harmonic pencil.

Proof. Suppose that l_1, l_2, l_3 and l_4 (see Fig. 110) are the sides of a complete quadrilateral. Let the vertices be denoted by P_{12}, P_{13}, P_{14}, P_{23}, P_{24}, P_{34} where P_{ij} is the point determined by the lines l_i

[1] It is implied in these definitions that on each side (vertex) of the complete quadrangle (quadrilateral) there are two and not more than two vertices (sides).

and l_j. The diagonals will be designated most conveniently by the self-explanatory symbols $d_{12,34}$, $d_{13,24}$ and $d_{14,23}$; finally a line through a vertex P_{ij} and through the point of intersection of the two diagonals which do not pass through P_{ij} will be denoted by d_{ij}. All these notations are illustrated in Fig. 110. Since there are six vertices, it looks as if the proof of the theorem would have to consist of six parts. For the vertex P_{12} we have to prove, using the notation of Definition LVIII*b* that $(l_2 d_{12} l_1 d_{12,34}) = -1$. The reader can verify without difficulty that the statement of the theorem *and its proof* for any other vertex P_{ij} can be obtained from the statement and proof for P_{12} by replacing the indices 1 and 2 by i and j respectively and adjusting the last two subscripts in $d_{12,34}$ so as to include all four numbers 1, 2, 3, 4 (compare *171*, 14). We can therefore content ourselves with the proof for this one case only; the notations $Q_{12,3}$ and $Q_{12,4}$ for the intersections of d_{12} with l_3 and l_4 and $R_{12,34}$ for the intersections of the two diagonals which do *not* pass through P_{12} and P_{34} complete the equipment needed in this case.

We have now (remember also the notation introduced in Definition LIX*b*, p. 404):

(15.9)

$$l_2 d_{12} l_1 d_{12,34} \overset{l_4}{\underset{\wedge}{=}} P_{24} Q_{12,4} P_{14} P_{34} \overset{P_{12}}{\underset{\wedge}{=}} P_{23},\ Q_{12,3} P_{13} P_{34} \overset{R_{12,34}}{\underset{\wedge}{=}} P_{14} Q_{12,4} P_{24} P_{34}.$$

Hence

$$(15.10)\qquad (P_{24} Q_{12,4} P_{14} P_{34}) = (P_{14} Q_{12,4} P_{24} P_{34}).$$

On the other hand, it follows from *168*, 13, that

$$(P_{14} Q_{12,4} P_{24} P_{34}) = (P_{34} P_{24} Q_{12,4} P_{14});$$

and from *168*, 12 that

$$(P_{24} Q_{12,4} P_{14} P_{34}) = \frac{1}{(P_{34} P_{24} Q_{12,4} P_{14})};$$

so that

$$(P_{14} Q_{12,4} P_{24} P_{34}) = \frac{1}{(P_{24} Q_{12,4} P_{14} P_{34})}.$$

This result, in combination with (15.10), shows that

$$(P_{24} Q_{12,4} P_{14} P_{34})^2 = 1, \quad \text{so that} \quad (P_{24} Q_{12,4} P_{14} P_{34}) = \pm 1.$$

The first of these possibilities is excluded[1] by means of the corollary to Theorem XCIII; hence $(P_{24} Q_{12,4} P_{14} P_{34}) = -1$. If we combine

[1] To make this corollary strictly applicable, it has to be shown that no two of the points P_{24}, $Q_{12,4}$, P_{14} and P_{34} can be coincident.

this result with the first perspectivity of (15.9), we conclude that $(l_2 d_{12} l_1 d_{12,34}) = -1$, as was to be proved.

There is an unmistakable quality about this theorem which can not help making an appeal to our æsthetic sense. It is perfectly general, in so far as it applies to *every* quadrilateral, it has nothing to do with measurement — and it has many interesting consequences.

Before discussing any of these consequences, let us consider the corresponding theorem for the complete quadrangle. Its statement is obtained by applying to Theorem XCVI the point ⟷ line correspondence, discussed at the end of *167*. This gives the following result:

Theorem XCVII. Two vertices of a complete quadrangle are harmonic conjugates with respect to the diagonal point on the side determined by them and the point of intersection of this side with the line joining the other two diagonal points; these four points form a harmonic pencil.

The reader should have no difficulty in proving this theorem; it can be done in various ways. One of them consists in following out, step for step, the proof of Theorem XCVI, and making in each step the change point ⟷ line. This will go through, because all the concepts and theorems which are used in the preceding proof are paralleled by concepts and theorems obtainable from them by this same change — compare (15.2) and (15.2*a*), Definitions LVII*a* and LVII*b*, LVIII*a* and LVIII*b*, LIX*a* and LIX*b*, LX, Theorems LXXXV, LXXXV*a*.

The procedure which we have here illustrated is applied a great deal throughout Projective Geometry.

It was explicitly recognized by the founder of Projective Geometry, Jean Louis Poncelet (1788–1867), and by the early workers in this field. But until comparatively recent times it was looked upon more as a "heuristic principle," i.e. as a means of discovering new theorems, than as a method of proof — although used extensively, its rôle and significance were not always clearly understood. In recent years, it has been completely justified on the basis of explicitly stated assumptions [1] and has thus become a valid argument in proof; it is known as the *principle of duality for projective geometry in the plane.*

Principle of Duality in the plane: With every theorem involving points and lines there is associated another such theorem obtained

[1] Compare, e.g., Veblen and Young, *Projective Geometry*, Vol. 1, pp. 26–29.

from the former by interchanging the words and phrases relating to *point* and *line*, and then making such linguistic modifications as good usage requires. A proof of either of these theorems constitutes also a proof of the other.

Of two such theorems each is called the *dual of the other*. It may happen that the dual of a theorem is equivalent in content to the theorem itself; it is then called a *self-dual theorem*.

The principle of duality exhibits a certain symmetry in the structure of projective geometry, which is of very great value in the exploration of this field. Theorems occur in pairs; for instance, with Theorem XCIV is associated the theorem referred to on page 413, viz.:

Theorem XCIVa. If three pairs of elements of two pencils of lines with different vertices are given, one and only one projectivity exists between these pencils, in which these pairs consist of corresponding elements.

Similarly there is a dual of Theorem XCV, viz.:

Theorem XCVa. A correspondence between two pencils of lines which leaves the double ratio of any four elements invariant is a projectivity.

Of the following theorem, known as the *theorem of Desargues*, and of fundamental importance in projective geometry, the dual is identical with the converse:

If the vertices of two triangles can be made to correspond in such a way that the lines joining corresponding vertices have one point in common, then the points of intersection of the sides opposite corresponding vertices lie on one line.

In the projective geometry of three-dimensional space there is also a principle of duality; it is based on a point $\longleftrightarrow$ plane interchange, the line being a self-dual element. The general concept of duality permeates other parts of mathematics.

Even though we have not done more than suggest the plausibility of the principle of duality in the plane we shall use it freely in the further developments.

171. In the midst of glory.

1. Show that, in Fig. 110, there is a projectivity between the points on l_1 and those on l_4 in which P_{12} and $Q_{12,4}$, P_{13} and P_{14}, $Q_{34,1}$ and P_{34}, P_{14} and P_{24} are corresponding pairs.

2. Show that the pencils of lines through P_{13} and P_{14} are projectively

related, if l_3 and l_1, $d_{13,24}$ and l_4, l_1 and $d_{14,23}$, d_{13} and $P_{14}Q_{13,2}$ are pairs of corresponding lines.

3. Show that there is a projective relation between the pencil of lines through P_{14} and itself, in which l_4 and $d_{14,23}$, $d_{14,23}$ and l_4, d_{14} and $P_{14}Q_{13,2}$ are pairs of corresponding lines and l_1 is a self-corresponding line.[1]

4. Show that there is a projective relation between the pencil of points on l_1 and itself, in which P_{13} and P_{12}, $Q_{34,1}$ and $Q_{24,1}$, P_{12} and P_{13} are pairs of corresponding points and P_{14} is a self-corresponding point.[1]

5. Show that the double ratio $(P_{12}Q_{34,1}P_{14}P_{13}) = 2$, while the double ratio $(P_{13}Q_{34,1}P_{14}P_{12}) = -1$.

6. Show that in the projectivity determined in 4, to the point $Q_{24,1}$ as belonging to the first pencil of points on l_1 corresponds the point $Q_{34,1}$ as belonging to the second.

7. Prove that if, in a projectivity between two pencils on the same basis or on the same vertex, there is given one self-corresponding element and one pair of doubly corresponding elements (as in 3 and 4), then any two corresponding elements which are distinct are doubly corresponding.

8. Devise a geometric construction for a point S, such that the double ratio of $(PQRS) = -1$, where P, Q and R are given collinear points. (Compare Theorems XCII and XCIII.)

9. Generalize the construction of 8 to the case in which $(PQRS)$ is to be equal to an arbitrary real number.

10. Prove that a projectivity of two pencils of lines is determined when three pairs of corresponding lines are given (compare Theorem XCIV*a* and Fig. 108).

11. On a straight line l three pairs of points, A_1 and A'_1, A_2 and A'_2, A_3 and A'_3 are given. Construct a chain of perspectivities connecting the sets $A_1A_2A_3$ and $A'_1A'_2A'_3$ so as to make these pairs consist of corresponding elements of a projectivity.

12. Carry out a construction similar to that of 11 for lines through a point.

13. Prove that a projectivity between a pencil of points on a line (lines on a point) and itself is uniquely determined if one self-corresponding element and two doubly corresponding elements are given.

Remark. Projectivities of this type are of special interest and have received the special title of *involutions;* examples are furnished by Exercises 3 and 4.

14. Prove Theorem XCVI for the point P_{13} and verify the assertion made on page 416, by showing that this proof can be obtained from the one given on that page by replacing the subscript 2 by the subscript 3.[2]

[1] Here it is to be understood that the first named lines (points) in each pair refer to elements of the pencil of lines (points) through P_{14} (on l_1) taken in one order and the second named lines (points) to elements of the same pencil taken in another order.

[2] The ambitious reader can add four additional exercises at this point!

15. Prove Theorem XCVII independently of the principle of duality (see p. 417; devise a suitable notation).

16. State the dual of the following theorems:

(a) If two lines have more than one point in common they are coincident.

(b) If the lines joining corresponding vertices of two complete quadrangles are concurrent, then the points of intersection of corresponding sides lie, three by three, on four lines.

(c) If the lines joining corresponding vertices of three triangles are concurrent, the three lines on which the corresponding sides of the triangles, taken two by two, meet are also concurrent.

17. Show that the 6 vertices of a complete quadrilateral can be combined in three different ways so as to form three different complete quadrangles; dualize this problem.

18. Make clear the relations between vertices, sides, diagonals and diagonal points of the complete quadrilateral and the complete quadrangles designated in 17.

19. Prove that on any diagonal of a complete quadrilateral the two vertices are harmonic conjugates with respect to the points in which the diagonal is met by the other two diagonals.

20. If two projective pencils of points are situated on two distinct lines, and if the point of intersection of these lines is a self-corresponding element, the projectivity is a perspectivity. Prove this statement.

21. Prove the dual of the theorem stated in 20.

172. A view of a familiar region. It must have become clear at this stage of the study of projective geometry that the possibilities for development which this subject presents are very extensive.

An adequate treatment would require an indication not only of the extent of the subject but also of a way in which the multifarious material can be systematically explored and brought under control. Neither of these things can be done here; they require the introduction of analytical methods, the study of algebraic forms, and many other topics of great interest. The books already referred to offer a starting point for such undertakings. In the few pages to which the rest of this chapter is confined, we can not do more than further whet the appetite by bringing forward some of the most fascinating among the more easily accessible theorems.

Theorem XCVIII. Two points, the midpoint of the segment joining them and the ideal point of the line determined by them constitute a harmonic pencil of points.

Proof. (See Fig. 111.) If A_2 is the midpoint of the segment A_1A_3, then $A_1A_2 = A_2A_3$ and hence $\frac{A_1A_2}{A_2A_3} = 1$. Moreover we know from *168*, 10 that if A designates the ideal point on the line a, then $\frac{A_1A}{AA_3} = -1$.

Consequently, $\frac{A_1A_2}{A_2A_3} \div \frac{A_1A}{AA_3} = -1$, as was to be proved (compare Definitions LVIIIa and LXI).

Corollary. Two sides of a triangle, the median through their common point and the line through this point parallel to the third side form a *harmonic pencil* of lines.

Any careful reader of this chapter can supply the proof.

The theorem and corollary show that if there is a way of distinguishing the ideal point on a line from every other point, then the midpoint M of a segment AB of the line can be defined as the

FIG. 111

harmonic conjugate of the ideal point with respect to the endpoints; and a line through P, parallel to the given line, as the line conjugate to PM with respect to PA and PB.

This fact illustrates a method by which, at least on a line, metric concepts, i.e. such as have to do with measurements, can be introduced into projective geometry, viz. by singling out a particular point upon the line. It is conceivable that point A^* on the segment A_1A_3, different from the ideal point, be singled out, and that the harmonic conjugate of A^* with respect to A_1 and A_3 be called the "midpoint" of the segment A_1A_3. It is one of the interesting aspects of projective geometry that it makes possible a generalization of the metric concepts with which ordinary experience has familiarized us. From this point of view the theorem as here stated represents a compromise between the metric and projective points of view; for it states a metric fact in terms of a projective concept. There are many such theorems, a natural result of the way in which projective geometry has been developed (compare p. 399). The familiar theorem that the diagonals of a parallelogram bisect each other is, in view of Theorem XCVIII, a special

case of *171*, 19. Textbooks on projective geometry[1] contain a great many theorems of this kind.

173. Through the underbrush to a summit. We have seen that, if two pencils of lines are in perspective, the points of intersection of corresponding lines of the pencil lie on one straight line. If the pencils are projectively related, this need no longer be the case. Indeed it will only be true if the projectivity is a *perspectivity*. In the general case the points determined by corresponding lines of the two pencils will lie on a curve. It is one of the interesting facts of projective geometry that such a curve is always a *conic section*, i.e. a curve obtained as the intersection of a plane with a circular cone. These curves are represented by equations of the second degree in x and y. Conversely, the locus of every such equation is a conic section, viz. an ellipse, a hyperbola (a pair of intersecting lines), or a parabola (a pair of parallel or coincident lines). The ellipse, hyperbola and parabola are the *non-degenerate conic sections;* the pair of coincident, intersecting and parallel lines are the corresponding *degenerate conic sections*. The proofs of these statements are discussed in analytical geometry.

In an analogous manner, we consider the lines joining corresponding points of two projective pencils of points. If the pencils are in perspective, these lines pass through a point. If they are not in perspective there exists a curve to which all these lines are tangent; the curve is called the *envelope* of the family of lines. As should be expected on the basis of the principle of duality, the envelopes of the lines joining corresponding points of two projective pencils of points are again conic sections.

The usual proof of the connection between projectively related pencils and conic sections proceeds according to the following outline.[2] The conics are defined projectively as follows:

Definition LXIVa. The locus of the points of intersection of corresponding lines of two non-concurrent projective pencils of lines, which are not in perspective, is called a *point conic*.

Definition LXIVb. The envelope of the lines joining corresponding points of two non-collinear projective pencils of points, which are not in perspective, is called a *line conic*.

[1] In addition to books mentioned before, we indicate the following textbooks on the subject in English: R. M. Winger, *Projective Geometry;* T. F. Holgate, *Projective Pure Geometry;* L. W. Dowling, *Projective Geometry;* J. W. Young, *Projective Geometry.*

[2] See, e.g., Veblen and Young, *Projective Geometry*, Vol. I, p. 109; Winger, *Projective Geometry*, p. 112.

The line conic is obviously the plane dual of the point conic.

On the basis of these definitions, it is then shown that the point conic and the line conic have the properties of conic sections as defined in analytical geometry, in particular that they can be represented by equations of the second degree in Cartesian coördinates; also, that a straight line does not meet a non-degenerate point conic in more than two points, and that from an arbitrary point there can not be drawn more than two tangents to a non-degenerate line conic. Moreover, it is shown that the vertices of the pencils of lines which determine the point conic lie on this conic and that the bases of the pencils of points which generate the line conic are themselves tangent to this conic.

There is a further fundamental property of conic sections which can be demonstrated readily by means of a little analytical geometry. But its proof calls for more technical knowledge of this subject than we have acquired. We shall therefore add this theorem to our long list of unproved propositions — it is like admitting a person to membership in a society without formal initiation upon the recommendation of members in good standing. This is the property in question:

(15.11) "Any five points, of which no three are collinear, determine one and only one non-degenerate point conic"; its dual is:

(15.11*a*) "Any five lines, of which no three are concurrent, determine one and only one non-degenerate line conic."

Having acquired all this knowledge, we can now *prove* a few further theorems which are among the most interesting in elementary projective geometry. In order to get access to them, we have made this hurried march with borrowed equipment.

Theorem XCIX. If the points on a non-degenerate point conic are joined to two different points on this same conic, the two pencils of lines so generated are projectively related.

Proof. In virtue of Theorem XCV*a* it will suffice to prove that the double ratio of any four elements of one pencil is equal to that of the four corresponding elements of the other. Let A_1, A_2, A_3 and A be four arbitrary points on the conic (see Fig. 112), which are joined to the points P and Q; let the joining lines be designated by p_1, p_2, p_3, p and q_1, q_2, q_3, q. We have to show then that $(p_1p_2p_3p)$

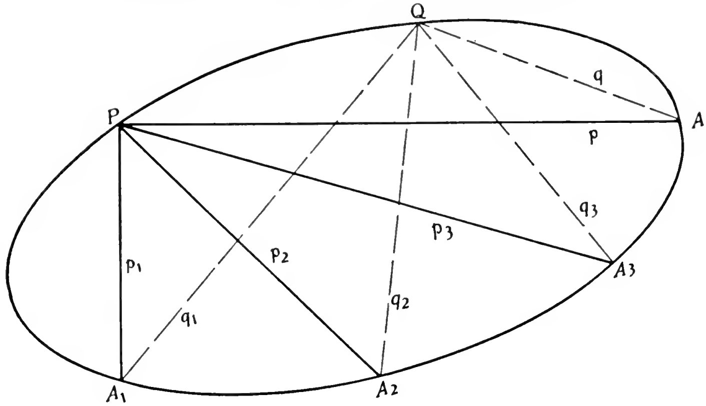


FIG. 112

$= (q_1q_2q_3q)$. The three pairs of corresponding elements p_1 and q_1, p_2 and q_2, p_3 and q_3 determine a unique projectivity (Theorem XCIV*a*). This projectivity associates with p a unique line through Q, let us call it q', such that $(p_1p_2p_3p) = (q_1q_2q_3q')$. Suppose that q' and q are different lines and that q' meets p in a point A'. Then the six points P, Q, A_1, A_2, A_3 and A' would lie on a point conic (see Definition LXIV*a*). But since the five points P, Q, A_1, A_2 and A_3 determine one and only one point conic (see (15.11)), this would be the given conic, so that the line p would have three points in common with this conic, viz. P, A and A'. But this would contradict one of the properties of conics, mentioned in preparation for this proof. Hence q' must coincide with q and thus our theorem is proved.

One point remains to be clarified. Among the lines of the pencil whose vertex is P, there is the line PQ; which line of the other pencil corresponds to it? The answer is that it must be the line which joins Q to Q, i.e. the tangent line to the conic at Q. Similarly, to the tangent line at P corresponds the line QP looked upon as an element of the pencil whose vertex is Q.[1]

Every statement in the proof of this theorem can be dualized so as to lead to the following dual theorem:

Theorem XCIXa. The intersections of the lines of a line conic

[1] It will be worth while for the reader to prove this statement by using the definition of the tangent line to a curve used in Chapter XII (compare Definition XLVI).

with any two fixed lines of this conic generate two projective pencils of points.

The form in which this theorem is frequently stated is the following:

The tangent lines of a conic intersect any two fixed tangent lines in two projective pencils of points.

The equivalence of this statement to the content of Theorem XCIX*a* rests on the facts (1) that the tangent lines of a point conic constitute a line conic, and (2) that the limiting positions of the points of intersection of the lines of a line conic generate a point conic. This relation between the two configurations is one of the significant elements of the projective theory of conics; it is another one of the many things which we have to leave for the interested reader to pursue.

Instead we shall go now to one of the strikingly beautiful theorems in this field.

Theorem C. If six points A_1, A_2, A_3, A_4, A_5 and A_6 be chosen arbitrarily on a point conic, then the points of intersection of the three pairs of lines A_1A_2 and A_4A_5, A_2A_3 and A_5A_6, A_3A_4 and A_6A_1 are collinear.

Proof. Let the point of intersection of the lines A_1A_2 and A_4A_5 be denoted by P, that of the lines A_2A_3 and A_5A_6 by Q, and that of the lines A_3A_4 and A_6A_1 by R (see Fig. 113); then we have to show that the points P, Q and R are collinear.

Consider the pencils of lines $A_1 - A_2A_3A_4A_6$ and $A_5 - A_2A_3A_4A_6$ [1]; from Theorem XCIX it follows that they are projectively related, i.e.

(15.12) $$A_1 - A_2A_3A_4A_6 \barwedge A_5 - A_2A_3A_4A_6.$$

The intersections of the lines of the first of these pencils with A_3A_4 are S, A_3, A_4, R; those of the second pencil with A_2A_3 are A_2, A_3, T, Q. Hence

$$A_1 - A_2A_3A_4A_6 \doublebarwedge SA_3A_4R \quad \text{and} \quad A_5 - A_2A_3A_4A_6 \doublebarwedge A_2A_3TQ.$$

From this fact, together with (15.12) and Definition LX, it follows that

(15.13) $$SA_3A_4R \barwedge A_2A_3TQ;$$

[1] The lines of those pencils have not been drawn in the figure, because they are not actually needed and would therefore unnecessarily complicate the diagram.

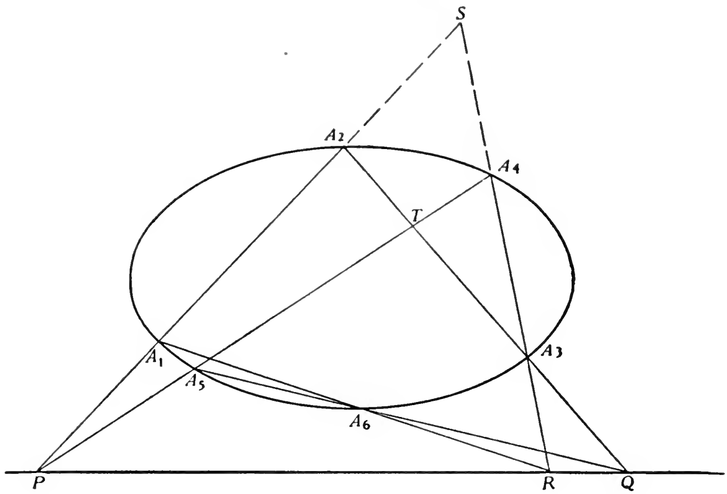


FIG. 113

i.e. that the points S, A_3, A_4 and R on the base A_3A_4, and the points A_2, A_3, T and Q on the base A_2A_3 are corresponding points of a projectivity, in which the point of intersection of the bases, viz. A_3, is evidently a self-corresponding point. From this we conclude, by means of *171*, 20, that (15.13) can be replaced by the stronger statement $SA_3A_4R \underset{\wedge}{=} A_2A_3TQ$, so that the lines SA_2, A_4T and RQ are concurrent. But SA_2 is the same line as A_1A_2 and A_4T the same as A_4A_5. Therefore, the intersection of the lines A_1A_2 and A_4A_5, i.e. the point P, must lie on the line RQ. But this is exactly the assertion of the theorem.

The converse of this theorem is true as well, viz.:

Theorem CI. If six points A_1, A_2, A_3, A_4, A_5, A_6, no three of which are collinear, are so situated, that the points of intersection of A_1A_2 and A_4A_5, A_2A_3 and A_5A_6, A_3A_4 and A_6A_1 are collinear, then these six points lie on a conic.

The importance of this result will be realized if compared with (15.11). Its proof will be left to the reader (compare *175*, 15). If six points on the conic are connected in any order whatever we obtain a hexagon, inscribed in the conic. Denoting them by A_1, . . . A_6, the line segments A_1A_2, A_2A_3, A_3A_4, A_4A_5, A_5A_6 and A_6A_1 are called the consecutive sides of the hexagon. Two sides

which in this sequence are separated by two consecutive sides are called *opposite sides;* thus A_1A_2 and A_4A_5, A_2A_3 and A_5A_6, A_3A_4 and A_6A_1 are the three pairs of opposite sides of the hexagon. With this terminology, the contents of Theorems C and CI can be stated as follows:

If a hexagon is inscribed in a conic, the points of intersection of the three pairs of opposite sides are collinear; and conversely.

In this form the theorem is usually referred to as Pascal's theorem on the "*hexagrammum mysticum.*" It was discovered, not as a theorem in projective geometry but as a theorem of the only geometry which had been conceived at that time, viz. of what is now called Euclidean metric geometry, by *Blaise Pascal* (1623–1662), when he was 16 years old.[1] This interesting circumstance has added a certain glamour to the theorem which, quite apart from such fortuitous enhancement, makes a strong claim upon the interest of intelligent human beings. This is not the place to inquire into the psychological basis for the impression this geometric theorem makes upon the human mind. Not to raise this question, but rather in order to provide more human beings with the opportunity of making its acquaintance, this theorem had to be included in our itinerary.

The dual of Pascal's theorem is of course also valid. It is known as Brianchon's theorem, after its discoverer *Charles Julien Brianchon* (1785–1864), who published it in 1806, as part of an extended treatise on curved surfaces.[2] We state it as follows:

Theorem Ca. If six lines of a line conic be chosen arbitrarily, $a_1, a_2, \ldots a_6$, then the lines joining the pairs of points a_1a_2 and a_4a_5, a_2a_3 and a_5a_6, a_3a_4 and a_6a_1 are concurrent [3] (see Fig. 114).

Its converse leads to our final theorem, viz.,

Theorem CIa. If six lines $a_1, a_2, \ldots a_6$ are so placed that the lines joining the points a_1a_2 and a_4a_5, a_2a_3 and a_5a_6, a_3a_4 and a_6a_1 are concurrent, then there exists a line conic to which these lines belong.

The validity of the last two theorems follows from Theorems C, CI and the principle of duality. Proofs independent of this principle, while therefore not necessary, are nevertheless instructive. In view of the observations made in the paragraph just preceding Theorem C, we can give to Brianchon's theorem the

[1] Compare, e.g., D. E. Smith, *Source Book in Mathematics*, pp. 326–330.

[2] Compare D. E. Smith, *op. cit.*, p. 331.

[3] The reader will doubtless understand that a_1a_2 designates the point of intersection of the lines a_1 and a_2, etc.; the present footnote is inserted for his reassurance.

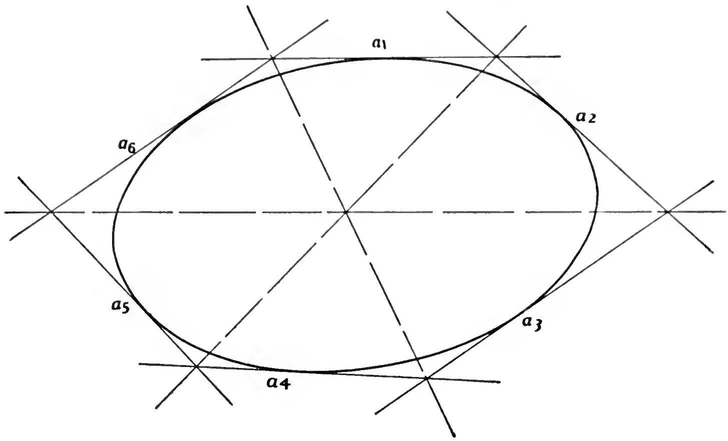


FIG. 114

following form, under which it frequently appears in books on metric geometry:

The lines joining opposite vertices of a hexagon circumscribed about a conic are concurrent.

174. Hints for the traveler. With the acquisitions of the preceding section, the purpose of this chapter has been accomplished. But it is unwise when we are in a museum, even if we have gone there for the express purpose of seeing a few particular paintings, not to take any notice of other treasures that may be stored there. Even if only for future use, a glance at a catalogue would be valuable. The analogy is far from perfect; for we shall not put before the reader any catalogue of theorems from projective geometry. The following section will give him the opportunity of testing the understanding and insight he has gained from the reading of the preceding ones. In the two remaining paragraphs we want to prepare the way for this venture by pointing out its connection with the rest of our expedition to the region of great fame.

It has already been observed that a good many theorems of projective geometry were known, long before this field was independently established, and that there are many statements which represent a compromise between metric and projective geometry. Thus when we make a distinction between a pair of parallel lines and a pair of "intersecting" lines, we are not true to the spirit of

projective geometry. Nevertheless, since we are also interested in metric geometry, these concessions have an unquestioned value. Several of the questions proposed in the next section are such special cases of theorems of projective geometry. The reader can simply be satisfied to recognize them as such; he can also amuse himself by looking for independent proofs.

Another source of some of the questions is the specialization which is obtained when of two points on a point-curve one tends toward the other, so that the line joining them becomes a line tangent to the curve (compare the footnote on p. 424); and, dually, when of two lines on a line-curve, one is made to tend towards the other, so that their point of intersection becomes a point of the curve. In such cases, the reader should not be satisfied with an independent proof of the theorems (indeed it is unnecessary!); he should also recognize them as special cases of more general theorems.

With this elucidation of its anatomy, the closing section of Chapter XV enters upon the scene.

175. A few short trips.

1. Derive the theorem that the diagonals of a parallelogram bisect each other from Theorem XCVIII.

2. Use Theorem XCVI and a straight edge (no other tools!) to construct the harmonic conjugate Q with respect to two given points A and B, of the arbitrary point P on the line determined by A and B (consider separately the case in which P is between A and B, and the case in which P is outside the segment AB).

3. Given two parallel lines. Construct the midpoint of a segment AB on one of them by means of a straight edge only.

4. Three points, A_1, A_2, A_3 are given arbitrarily on a line a, and three points B_1, B_2, B_3 on a line b, such that the lines A_1B_1, A_2B_2 and A_3B_3 are concurrent. Deduce from the theorem of Desargues (see p. 418), (1) that the points of intersection of A_1B_2 and A_2B_1, of A_2B_3 and A_3B_2 and of a and b are collinear: (2) that the points of intersection of A_1B_2 and A_2B_1, of A_1B_3 and A_3B_1, and of a and b are collinear, (3) that the four different points mentioned in (1) and (2) are collinear. Show that the proof is valid whether or not the point of concurrence of A_1B_1, A_2B_2 and A_3B_3 lies between the lines a and b.

5. Derive the results stated in 4 from Theorem XCVI.

6. Given the lines a and b, and a point P not on either of them. Construct a line through P and through the intersection of a and b, without using the latter point.

7. Carry out the construction of 6 in case a and b are parallel, thus obtaining through P a line parallel to two given parallel lines.

8. The points A_1, A_2, A_3, A_4 and A_5 are taken arbitrarily on a point conic. Prove that the points of intersection of A_1A_2 and A_4A_5, of A_2A_3 and A_5A_1, of A_3A_4 and the tangent to the conic at A_1, are collinear.

9. Construct the line tangent to the conic at an arbitrary point.

10. The points A_1, A_2, A_3 and A_4 are taken arbitrarily on a point conic. Show that the points of intersection of A_1A_2 and A_3A_4, of A_2A_3 and A_4A_1, of the tangents at A_1 and A_3 are collinear.

11. A triangle is inscribed in a conic. Prove that the three points of intersection of the sides with the tangents at the opposite vertices are collinear.

12. State and prove the dual of the theorem in 8.

13. Similarly for 10.

14. Similarly for 11.

15. Given 5 points. Construct the second point in which the conic determined by them meets an arbitrary line through one of them.

16. Given 5 lines. Construct the second line through an arbitrary point on one of them, which is tangent to the conic determined by them.

17. Prove Theorem CI (see p. 426).

CHAPTER XVI

A FINAL LOOK AT THE MAP

The human interest in exploration in any field lies in the log book of the journey, the difficulties met, and how they were overcome, until the object of the expedition is attained. . . . Human aspiration and the spirit of conquest are the main motive power of all such endeavors, and not the practical advantages which accrue from the results achieved. So should the objects of pure research in the realm of pure science be regarded, and from this point of view should the thoughts and experiences of the adventure be described. — Sir Richard Gregory.

176. For the future. The time has come for the author to take leave of his patient readers. He can not know how many who accompanied him on the early journeys through some of the regions of mathematical science have persisted to the end, through thorny brambles relieved from time to time by vistas from high places. Some among them have probably waited outside one of the labyrinths to rejoin the party when there was promise of a more immediate reward for their labors. If they have chosen the resting places with good judgement, no harm has been done. It is not necessary to see all the sights in a foreign country to get some idea of its charms. Indeed this journey by no means pretends to lead to all the beautiful spots of the region it traverses, not even to all those which are accessible to the inexperienced traveler. There are a good many which are not included in our itinerary. Before parting and returning to their normal occupations, the participants in this expedition should hear about one of these; it will provide them with an incentive for future adventures. Perhaps this account should have come more appropriately before we had reached the end. No opportunity has been found for it at an earlier time; it may prove to be a fitting means for bringing the book to a close.

177. A geography lesson. When we make a map of a part or the whole of the earth's surface, we usually wish to have any two of its sections which have a common boundary, distinguished from each other by different colors. On a map of the United States for instance, Illinois and Indiana, Ohio and New York, Alabama and

Mississippi, Washington and Oregon should be differently colored. We do not object however if California and Vermont have the same color. Similarly, we do not mind it if on a map of the state of Michigan, the colors of Clare County and Midland County are identical.[1] But we would not want to admit the Counties of Weld and Arapahoe in Colorado with the same tint. The following question now arises: How many different colors are necessary and how many are sufficient to color any map?

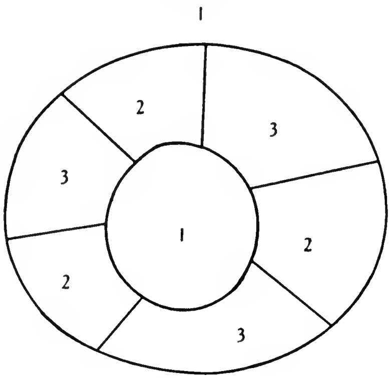


FIG. 115

To get an idea of the problem, we shall begin by considering some maps of our own invention. For the one represented in Fig. 115, three colors will obviously suffice. If it were an island, so that the part of the plane surrounding it also requires a color, even then we could manage with three colors. This is no longer true for the map in Fig. 116; but we could take care of this with four colors, also if there is to be a sea on the outside.

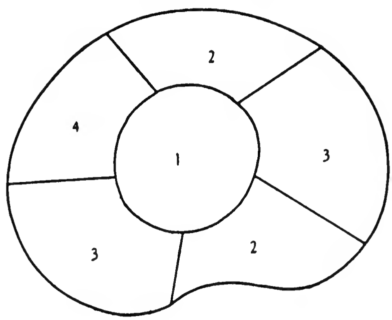


FIG. 116

On the other hand, a map such as the one shown in Fig. 117*a*, appears to need six colors. But, if the hint intended in the footnote on this page has taken effect, we will recognize that this would be undue extravagance; as shown in *b* and *c*, for such a map, two or at most three colors would suffice. Two sections of a map which, like idealized cuts of a dis-

[1] Are we assuming a knowledge of the geography of the state of Michigan too detailed for the inhabitants of Utah or of South Carolina? Warren and Asper Counties in Iowa, Columbia and Hempstead Counties in Arkansas, or Frio and Macmullen Counties in Texas would serve our purpose just as well. But there are not many suitable examples of this kind in the New England states.

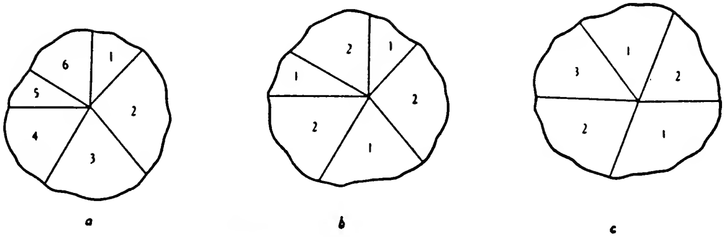


FIG. 117

torted pie, have only a single point in common, do not need to be colored in different ways. Even if they have two points in common as some of the regions in Fig. 118, they may carry the same color.

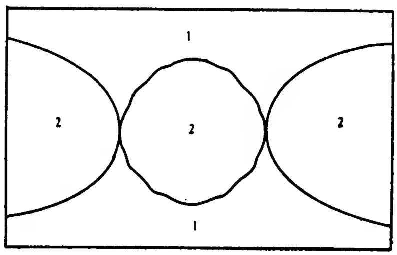


FIG. 118

If we accept this principle, our problem can be stated somewhat more precisely as follows:

> How many colors are necessary and how many are sufficient to color any map, if we require that any two of its regions with a common boundary shall have different colors, it being understood that a boundary must consist at least of a segment of a curve and not of isolated points?

The examples which we have been examining indicate that *at least four colors are necessary.* Are they also sufficient? This simple question is as yet awaiting an answer. Raised for the first time in 1878 by the famous British mathematician Arthur Cayley (see p. 279), it turned out to be a much more difficult problem than any one would have imagined. Many papers have been written on the subject; in recent times a number of American mathematicians have contributed to its discussion, but — it is still unsolved. It has been shown that *five colors are* sufficient.[1] Moreover no one

[1] A simple proof of this fact can be found in the book by H. Rademacher and O. Toeplitz, *Von Zahlen und Figuren*, pp. 54–62, (see p. 148). Accounts of the history and development of the "four-color problem" are found in an article by H. Brahana, *American Mathematical Monthly*, Vol. 30, 1923, p. 234, and in a paper by A. Errera, *Periodico di Matematiche*, Series IV, Vol. VII, 1927, p. 20.

has yet succeeded in constructing a map for which more than four colors are needed. This is the position at the present time of this fascinating and puzzling question. It makes a good starting point for another journey, into the realm of mathematics that is called *topology*, or *analysis situs*.

178. The end of the first day. For us however it must be the terminal point. Indeed it was deliberately reserved for this use for two reasons.

In the first place, in order to emphasize the facts that mathematics was not "finished 500 years ago" as seems to be believed by a great many people, even among the "educated classes"; that mathematics is not merely concerned with numbers; that it is a growing science, forever young; that it contains many unsolved problems and unanswered questions; that it does not require unusual intelligence to know what some of these questions are.

In the second place, because it is as a map that this book was conceived. The analogy of traveling through a new country which has been used, although not consistently, at many points, was not meant only as a convenient figurative way to lead us on when courage failed, but also as an indication of the pattern on which we have worked. To guide those who are not specialists nor intend to become specialists through the extraordinary domain of human knowledge and invention which is called mathematics, to invite them to an understanding of what it is concerned with, to orient them in this field which comprises in abstract form so many and such diverse aspects of human experience, leaving much for further exploration and discovery, this has been the purpose for which the writing of this book was undertaken.

INDEXES

LIST OF DEFINITIONS

indicating the page on which they are stated.

LIST OF NUMBERED EQUATIONS

indicating the page on which they occur for the first time. (The digits to the left of the decimal point indicate the chapter.)

LIST OF FIGURES

indicating the page on which they occur.

LIST OF SYMBOLS AND ABBREVIATIONS

indicating pages on which they are explained.

LIST OF THEOREMS

indicating the page on which each theorem is stated.

INDEX

The numbers refer to pages; a number in italics refers to a footnote on the corresponding page.

A

B